ELSEVIER

Design Paper

Design of a Group-Randomized *Streptococcus pneumoniae* Vaccine Trial

Lawrence H. Moulton, PhD, Katherine L. O'Brien, MD, MPH, Robert Kohberger, PhD, Ih Chang, PhD, Raymond Reid, MD, MPH, Robert Weatherholtz, BS, Jill G. Hackell, MD, George R. Siber, MD, and Mathuram Santosham, MD, MPH

The Center for American Indian and Alaskan Native Health, Department of International Health, The Johns Hopkins University School of Hygiene and Public Health, Baltimore, Maryland (L.H.M., K.L.O., R.R., R.W., M.S.), Department of Biostatistics, The Johns Hopkins University School of Hygiene and Public Health, Baltimore, Maryland (L.H.M.), Wyeth Lederle Vaccines, Pearl River, New York (R.K., I.C., J.G.H., G.R.S.), and Department of Pediatrics, The Johns Hopkins University School of Medicine, Baltimore, Maryland (K.L.O., M.S.)

ABSTRACT: A group-randomized, double-masked, phase III trial of a *Streptococcus pneumoniae* conjugate vaccine is being conducted in American Indian populations in the southwestern United States. Approximately 9000 infants will be enrolled in the primary efficacy cohort with vaccine allocation determined by community of residence. The trial is designed to continue until 48 cases of invasive pneumococcal disease due to vaccine serotypes have accumulated. Thirty-eight geographically and socially distinct areas were randomized within blocks formed by population size and geographic location. This design affords the opportunity to capture the effects of herd immunity (indirect effects) by estimating the impact of the vaccine intervention on nonimmunized infants. Group-randomized trials have challenging design and analysis features, many of which are discussed here in the context of the first such trial designed to lead to licensure of a drug or biologic in the United States. *Control Clin Trials* 2001; 22:438–452 © 2001 Elsevier Science Inc. 2001

KEY WORDS: *Vaccine, group-randomized trial*, Streptococcus pneumoniae

Address reprint requests to: Lawrence H. Moulton, PhD, Department of International Health, The Johns Hopkins University School of Hygiene and Public Health, 615 North Wolfe Street, Baltimore, MD 21205 (lmoulton@jhsph.edu).

Received August 21, 2000; accepted February 16, 2001.

Controlled Clinical Trials 22:438–452 (2001)

655 Avenue of the Americas, New York, NY 10010

0197-2456/01/$–see front matter
PII S0197-2456(01)00132-5

INTRODUCTION

Invasive pneumococcal disease is a major cause of mortality and serious morbidity in children throughout the world. Incidence rates for children under 2 years of age are estimated to be 100–160 per 100,000 person-years in the United States [1, 2] and many times greater in the developing world [3]. In a recent five-country study, 20% of the serious bacterial infections in infants less than 4 months of age were caused by pneumococcus [4]. Infection with *Streptococcus pneumoniae* is the leading cause of pneumonia in children under 5 years of age [5]. Of the 4,000,000 children and infants who die each year throughout the world, approximately 20–40% of the deaths are estimated to occur as a result of pneumococcal pneumonia [6]. In the United States, *S. pneumoniae* is the leading cause of community-acquired bacterial pneumonia, and in all age groups causes about 40,000 deaths annually [7]. In addition, rates of penicillin resistance are rising, and in some parts of the United States over 50% of pneumococcal isolates are nonsusceptible to one or more antibiotics [8]. The currently licensed 23-valent pneumococcal polysaccharide vaccines are poorly immunogenic for important serotypes in infants and young children. Thus, the importance of developing pneumococcal vaccines to prevent infection among young children is great, and several vaccine manufacturers have products in various stages of development.

We have designed and implemented a group-randomized vaccine trial to estimate the total efficacy with respect to pneumococcal invasive disease of a heptavalent *S. pneumoniae* protein conjugate vaccine in Native American populations (M. Santosham, Principal Investigator). The incidence of pneumococcal diseases among the American Indian and Alaska Native populations has been documented at four to ten times that of the general U.S. populations [9–12]. The Johns Hopkins University Center for American Indian and Alaskan Native Health has a long history of collaborating with these communities in efforts to reduce their infectious disease burden, from the introduction of oral rehydration protocols to the conduct of *Haemophilus influenzae* type b (Hib) conjugate vaccine trials.

When we developed our trial design using a seven-valent (serotypes 4, 6B, 9V, 14, 18C, 19F, 23F) pneumococcal protein conjugate vaccine (Wyeth Lederle Vaccines [WLV]), there was an ongoing trial of the same product in a population of 38,000 infants in northern California [13]. At the time, the number of cases of invasive pneumococcal cases that had accrued was small, and there was some concern about whether the attack rates would be as high as had been projected. In that trial, vaccine was allocated using individual randomization with the goal of yielding the classic estimate of vaccine efficacy, namely the degree to which an individual is protected following immunization. The group-level allocation scheme we adopted enables us to estimate the total efficacy, which is a function of both individual-level protection and protection afforded through reduction in secondary attack rates and transmissibility, and mimics the effect of vaccine on a community in a postlicensure era.

Our trial was designed with two goals in mind: (1) to evaluate nonstandard but useful aspects of vaccine effectiveness, in particular, the total efficacy and indirect effects of the vaccine when introduced on a mass scale in a closed population; and (2) to serve as a pivotal trial to demonstrate efficacy. Thus, the study was designed to serve as a pivotal trial leading to Food and Drug Administration (FDA) licensure of the vaccine in the U.S. pediatric population and to determine the effect of the vaccine on disease in the total population, including those unimmunized (i.e., indirect effects).

Ours is the first group-randomized trial conducted in the United States designed to lead to product licensure. The following sections, besides providing details on the trial design, explain the history of the trial design and its rationale. Special attention is given to statistical issues that are unique to such designs and to our attempts to address them. The review article by Hayes et al. [14] contains a more general discussion of group-randomized trials of infectious disease interventions.

DESIGN

Vaccines

The heptavalent pneumococcal protein conjugate vaccine (7VPnC) used in this study is manufactured by WLV. It contains saccharides of seven pneumococcal serotypes (4, 6B, 9V, 14, 18C, 19F, 23F) conjugated by reductive amination to a carrier protein (CRM_{197}, a nontoxic variant of diphtheria toxin isolated from cultures of *Corynebacterium diphtheria*).

In the control arm of the trial we employed an investigational meningococcal C conjugate vaccine (MnCC). This vaccine, not yet licensed for use in the United States (although licensed during our trial in October 1999 in the United Kingdom), also is manufactured by WLV, using the same carrier protein (diphtheria CRM_{197}). An outbreak of meningococcal C disease occurred among children on the Navajo Nation in 1995, and we thought this vaccine might provide benefit to children living in the control arm areas.

We wanted to conduct a trial that would parallel as closely as possible a postlicensure situation in which most young children are immunized, thereby maximizing the potential for manifestation of indirect effects. All children under 2 years of age, therefore, were eligible for enrollment.

For each of the study and control vaccines, three immunization schedules were designed according to age of entry into the trial: 6 weeks to 6 months (three doses, ideally at 2, 4, and 6 months of age, and a booster at 12–15 months of age), 7 months to 11 months (two doses 1 month apart, and a booster at 12–15 months of age), and 12 months to 23 months (two doses separated by at least 2 months). Over the course of the trial, the great majority of new enrollees are in the first group, which we refer to as the primary efficacy cohort.

Parameters of Interest

We performed a group randomization to enable the estimation of endpoints suitable for regulatory licensure, as well as to yield more detailed descriptions of the vaccine effects when introduced into a population on a mass scale. It has taken many years following the introduction of Hib conjugate vaccine to obtain estimates of the degree of extra population-level protection afforded by indirect (herd) effects, and it is still a matter of some controversy [15–17]. Designing direct estimation of these effects into a pivotal trial can provide information that will be useful for cost-effectiveness analyses that will have to be performed on a country-by-country basis.

We have adopted the definitions and design classifications described by Halloran et al. [18, 19] and applied them to our setting. Standard, individually randomized, phase III vaccine trials are designed to estimate the direct effect of

immunization (i.e., the benefit that accrues to the individual as a function of becoming protected against challenge by an infectious agent). In such a design, individuals in both arms of the trial are assumed to be mixing independently with each other in the community. Our group-randomized study design is depicted in Figure 1. "Units" are geographically distinct areas to which either study or control vaccine is randomized. Eligible, consenting individuals within any unit are all administered study or control vaccine, depending on the treatment arm to which their unit was allocated. Many factors may determine which individuals within a given unit decide to participate in the trial. Some of these factors may be related to risk of disease (or more accurately, risk of becoming diagnosed as a case), including socioeconomic status, distance to the nearest health facility, number of siblings, and day-care utilization. Thus, in our trial, comparison of attack rates between study participants (vaccinated) and nonstudy participants (unvaccinated) within 7VPnC vaccine units, which estimates direct effects, may result in a substantially biased estimate of vaccine efficacy. For example, if nonstudy participants have a much higher underlying risk of disease, then the efficacy of the vaccine will be biased upward.

Comparing the disease incidence among study participants in the vaccine units with study participants in the control units yields an estimate of the total effect of immunization. It includes both direct and indirect effects of the immunization program. We expect that immunizing the majority of infants and toddlers in the vaccine units will reduce the secondary attack rate through reduction of the number of infected individuals and of the number of those who carry *S. pneumoniae* in the nasopharynx. The result would be a synergistic effect, commonly referred to as the effect of herd immunity [20]. The sum of all the benefits to the participants is referred to as the total effect. The overall effect of immunization is estimated by comparing attack rates among all residents of the vaccine units with all residents of the control units, recognizing that only some fraction of the residents in each type of community actually received study or control vaccine. Because the overall effect is so dependent on the attained coverage (proportion enrolled), it is perhaps less generalizable than the other effect measures. However, it will permit some extrapolation to other populations with similar expected coverage levels. For example, if two thirds

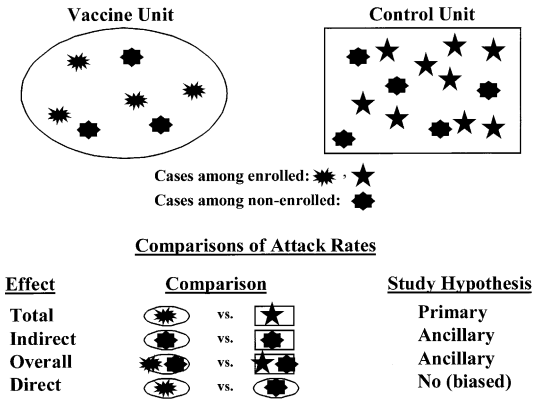

Figure 1 Schematic of trial study hypotheses according to the nomenclature of Halloran et al. [19]. Participants in each vaccine unit receive 7VPnC vaccine, while those in each control unit receive MnCC vaccine.

of infants are immunized, and there is 90% less disease overall in the vaccine units than in the control units, this will be important information for health systems in less developed areas of the world where lower vaccine coverage is common. It will be of less long-term interest for our study populations and other populations in the United States that, following licensure, are expected to attain over 90% coverage.

Analysis of the pneumococcal disease experience of those who are not enrolled in the trial will enable us to obtain an estimate of the indirect (herd) effect of vaccine. The indirect effect is estimated by the degree to which the incidence rate in nonparticipants in the vaccine units is lower than the rate in nonparticipants in the control units. Note that because the comparison is across sets of randomized units, this estimator will not suffer from the bias problem that we saw for the within-unit estimator of direct effect.

Objectives

1. To demonstrate that the 7VPnC vaccine is effective in preventing vaccine-type invasive pneumococcal disease occurring in infants and toddlers who have enrolled by 6 months of age and have completed the recommended primary three-dose infant series, or completed the primary three-dose infant series and the scheduled toddler booster, according to age- and window-appropriate schedules.

More specifically, the time at risk for invasive pneumococcal disease episodes is from 14 days after the last of the primary series until the end of the study or, if no booster has been received, until 16 months of age. This objective is to evaluate strict per-protocol efficacy, as is typical for vaccine trial study designs [21]. According to the nomenclature of Figure 1, this estimates the total 7VPnC effect among adequately immunized participants who fall in the early enrollment group (less than 7 months of age), which corresponds to how the vaccine would be employed in the general population following licensure. The evaluation is made by comparing disease incidence following the primary immunization series in those study participants age-appropriately vaccinated with 7VPnC to those age-appropriately vaccinated with MnCC.

2. To determine whether the 7VPnC vaccine prevents vaccine-type invasive pneumococcal disease in all children enrolled in the study.

This estimates the total effect on an intent-to-treat basis. The evaluation is made by comparing disease incidence in those study participants who received at least one dose of 7VPnC to those who received at least one dose of MnCC. The main version of this objective will analyze only those who enroll before 7 months of age to correspond to the primary efficacy analysis. A secondary version will include the older cohorts. We note that this differs from the more commonly defined intent-to-treat analysis that is based on all who have been randomized; such an approach would include all residents in the designated study areas, regardless of enrollment status. That analysis is described next in 3(a), but is of less interest because of its heavy dependence on enrollment levels.

3. To evaluate the effectiveness of the 7VPnC vaccine in preventing: (a) vaccine-type invasive pneumococcal disease in all children less than 2 years of age, regardless of enrollment in the study; and (b) all invasive pneu-

mococcal disease, regardless of serotype, in all children less than 2 years of age, regardless of enrollment in the study.

These evaluate overall effects of the immunization campaign.

4. To evaluate the impact of the 7VPnC vaccine on colonization of the nasopharynx: (a) with vaccine-type pneumococci; and (b) with any pneumococci regardless of serotype.

These evaluations are being performed in a subset of the main study populations (K. O'Brien, Naso-Pharyngeal [NP] Substudy Principal Investigator). Among other objectives, they will provide the biological basis for inference on reduction of pneumococcal vaccine serotype transmissibility and the associated indirect effect of immunization. In this nested study vaccinated children and all other children younger than 6 years of age residing in the household are evaluated for pneumococcal NP colonization. NP swabs are collected 1 month following the primary vaccination series, at the time of the booster, and at 6 months following the booster dose. If any child within the household is carrying pneumococcus, two follow-up visits are scheduled at 1 and 3 months following the initial visits. NP swabs are tested for the presence of pneumococcus, and isolates are serotyped using a novel immunoblot method that will enable the detection of all serotypes that are present [22]. The group-randomized nature of this trial enables this study to be conducted without problematic mixing of recipients and nonrecipients of 7VPnC vaccine.

5. To evaluate the impact of 7VPnC vaccine on the rates of pneumococcal disease and colonization in nonstudy participants: (a) vaccine-type pneumococcal disease/colonization occurring in unimmunized children < 2 years of age; (b) vaccine-type pneumococcal disease/colonization occurring in unimmunized individuals of any age; and (c) vaccine-type pneumococcal disease/colonization occurring in unimmunized siblings of immunized children.

These can be used to evaluate the indirect effects of the 7VPnC vaccine. The focus will be on (a), the group most likely to experience a reduction in disease incidence. The same analysis will be done for children < 6 years old as well, the interpretation of which will be aided by the NP colonization results described in objective 4. Incidence rates in nonstudy participants in 7VPnC areas are compared to rates in nonstudy participants in MnCC areas.

Clinical Endpoint

We defined a case of invasive pneumococcal disease as the identification of pneumococcus from a normally sterile site in an eligible subject, where eligibility is defined by the specific analysis under consideration. Invasive cases were categorized as being either a vaccine serotype or a nonvaccine serotype depending on whether the identified serotype was one contained in the 7VPnC vaccine.

Study Populations

The study is being conducted on the White Mountain Apache Reservation (Arizona) and the Navajo Nation (Arizona, New Mexico, and Utah). The annual birth cohorts are approximately 360 and 4500, respectively. In addition to

the high rates of pneumococcal disease, a major advantage of these populations is the relatively low degree of migration to and from the study areas. For administrative and programmatic purposes, the Indian Health Service (IHS) has created catchment areas called service units. The great majority of IHS service unit residents receive all of their health care from IHS facilities.

Surveillance

All laboratories of hospitals on the reservations, and of hospitals adjacent to the reservations, are contacted on a daily or weekly basis to obtain reports of any invasive pneumococcal cultures obtained from Native Americans. Physicians have been encouraged to obtain blood cultures on all infants < 2 years of age who present with a temperature above 103° Fahrenheit, as well as for other specific suspected diagnoses, in accordance with recommended practice. Subisolates of all positive pneumococcal cultures from normally sterile body fluids are sent to a referral laboratory for confirmation and serotyping. This procedure is performed for all positive cultures, regardless of age of patient or whether the patient is a trial participant. As noted above, information on the nontrial participants contributes primarily to estimation of indirect effects of the vaccine.

Randomization Units

Several vaccine allocation designs were considered with varying levels of geographically defined randomization (allocation) units. In the early stages of planning, one proposal had been to provide 7VPnC vaccine to half of the Navajo Nation based on a geographic separation into two parts. This approach is common to demonstration projects or quasi-experimental designs. It was dropped from consideration, however, because of the threat of geographic confounding of the disease-vaccine relationship and the concomitant lack of a strong basis for statistical inference.

Also considered was a design with eight geographically determined randomization units, which would have used the eight IHS service units covering the Navajo Nation, with four in each vaccine group. If the rates of invasive pneumococcal disease had been many times greater than the historical data indicated, this could have been an acceptable design: higher rates would translate into greater relative precision and coupled with the anticipated high efficacy could furnish acceptable power. In addition, there would have been a minimum amount of "contamination," or mixing across units. Yet with only a handful of events expected in each of the eight units, and the much greater between-unit variability that accompanies larger units (see the section below on power), we would not have had sufficient power.

An alternative was to randomize at the chapter level. The Navajo Nation is divided into 110 geopolitical units called chapters, each of which has a major community with a chapter house where many social and community functions take place. This would have been a feasible solution, but the concern was that there might be substantial mixing of groups of people across chapter lines, especially for Head Start center and elementary school attendance. Too much cross-unit social interaction would diminish the indirect effects of the vaccine, which we wanted to be able to capture in our design. After further consulta-

tions with persons familiar with the social demography of the area, we decided to make randomization units composed of two to four chapters, each unit with an average of about 200 resident eligible infants and toddlers.

In January 1997 we convened representatives from many groups knowledgeable about the social and geographic aspects of the Navajo Nation. These included representatives from the Navajo Nation, the IHS, the Bureau of Indian Affairs, and current and former Johns Hopkins University-paid employees engaged as field staff for other studies. Many factors were considered, including location of Head Start programs, elementary schools, towns, roads, and topographical features. With the goal of combining chapters into randomization units so as to minimize social interaction between the units, initial randomization units were decided upon. In the following month, they were refined slightly by the Navajo Nation demographer (Larry Rodgers) to make 41 intended randomization units. The most pertinent feature of these units is that no Head Start program or elementary school had a catchment area that extended beyond the unit in which it was located. A few of these randomization units were merged to achieve population and geographic balance, as described in the next section.

The Navajo Area IHS service units extend slightly beyond the borders of the Navajo Nation and include Navajo who live both in rural environments and in a number of "border towns," but who do not reside in any given chapter. Native American residents of these areas were eligible for study inclusion. These areas were split up by placing each in the randomization unit that contained the geographically closest chapter.

The second major group of participants in the trial were members of the White Mountain Apache Tribe living on the White Mountain Apache Reservation 100 miles south of the Navajo Nation. They are expected to comprise approximately 14% of the total trial outcomes. In this area, which is centered in the town of Whiteriver, there were no useful political boundaries to rely upon for constructing residence-based randomization units. For simplicity, and to minimize contamination through mixing of the population across 7VPnC and MnCC areas, we split the area into two randomization units, essentially the town of Whiteriver versus the rest, with a corresponding approximately equal split on population. The total number of available randomization units for the trial was thus $41 + 2 = 43$.

Place of residence is defined as wherever the infant's nuclear family spent more than 6 months of the year before enrollment. For persons who migrated, either temporarily or permanently, within the study area, we decided to continue giving the infant the same vaccine with which the infant began the study, even if the infant moved between 7VPnC and MnCC areas. The rationale was to ensure the infant received a complete series of one of the vaccines, rather than an incomplete series of both. Although this decision has the potential to slightly decrease the indirect effects, it was deemed to confer a significant advantage to the individual infant. We expected this situation to occur for only 1 or 2% of the study participants.

Labeling and Masking of Treatment Assignment

We decided to assign six labels to the vaccines (B, F, H, M, T, U), with three labels for 7VPnC and three for MnCC. The grouping of these codes is known only to a statistician employed by the manufacturer (I. Chang) who has no other

responsibilities with respect to the trial other than handling treatment allocation and randomization issues. We decided not to use individually coded vials because an individual trial participant might seek health-care services from multiple provider sites. The probability of arriving at the correct grouping of the vaccine labels by a random guess is 0.05 (there are 20 combinations of three of the six labels), and with less than two expected cases of invasive disease per randomization unit, sufficient accumulation of data by study personnel, either consciously or unconsciously, was deemed highly unlikely. In addition, field staff were blinded as to serotype of the invasive disease cases, and thus did not know which ones would be likely to be prevented by an effective vaccine.

In the Navajo study area, after randomization associated each randomization unit with a treatment group, a single letter code was associated with each chapter in a randomization unit in such a manner to ensure that there would be at least two letter codes in every randomization unit. The grouping of the chapters into randomization units is not known to any of the field staff; thus, there are two levels of masking, one achieved through labeling, the other through disguising the randomization unit boundaries. In fact, in the listings of labels appropriate for each geographic area that field staff rely upon for vaccine allocation, local community names are often given instead of chapter names for those places where there is more than one community in a chapter, but the communities do not overlap chapters. Figure 2 is a mock-up of two randomization units and the potential levels of their subdivision.

Stratification and Randomization

Even with a relatively large number of randomization units (43), the potential for a randomization resulting in geographic or population imbalances is substantial. Our plan was designed to achieve an approximate balance between the two treatment arms with respect to the number of potential participants within each service unit, and the overall number of potential participants, based on 1990 U.S. Census data. We thought there might be a sizable random effect associated with each service unit's principal source of health-care services, with possibly differential detection rates for the study outcomes. Within

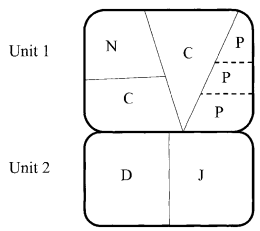

Figure 2 Subdivision and labeling of treatment randomization units. Hypothetical example: labels C, N, P are for one treatment; D, J, Z (Z does not appear here) for the other. Heavy solid lines mark the randomization unit boundaries; light solid lines mark the chapter boundaries; dashed lines mark local community boundaries. Unit 1 is composed of four chapters, one of which is further subdivided into three local communities. Unit 2 is composed of two chapters and has only two of the three possible treatment labels represented. Study personnel are unaware of the randomization unit bound.

each service unit, the strategy was to rank and then block randomization units by population size in blocks of two or four. Because of variability in the population sizes of the randomization units and the fact that some randomization units straddled two service units, we had to further restrict the randomization. Five of the unit assignments were restricted to be the same assignment as one or more other units in their blocks to prevent a high probability of extreme imbalance within service units. Thus, although we began with 43 identified units available for randomization, the actual randomization was carried out using only 38 independent randomization units, stratified using three blocks of four units and 13 blocks of two units.

Prior to randomization, we performed a test of the scheme by performing 498 randomizations and observing the chance of imbalance in population size, both overall and within each service unit. Overall, there was a 99.5% chance of having a better (closer to 50:50) than 45:55 (or 55:45) split. Within each service unit, there was a greater than 75% chance of better balance than 41:59, except for two service units whose 75th percentiles were 34:66 and 18:82 (the latter had a median of 32:68).

We considered performing a much more restricted randomization, one that would, say, randomly select one allocation from among the ten best potential allocations or execute randomizations until one meeting a specified criterion was reached. These would have guaranteed a given degree of balance, but could have opened the trial to criticism due to an apparent lack of complete randomization.

CORRELATION, SAMPLE SIZE, AND ANALYSIS ISSUES

Power and Length of Study

The trial was designed to continue until 48 primary efficacy cases of invasive pneumococcal disease caused by vaccine serotypes had accrued. This number was arrived at through the following considerations. First, for an individually randomized trial, 40 cases is the minimum number necessary to achieve at least 80% power for the lower bound of a 95% confidence interval to lie above 20% efficacy, if in fact the true efficacy is 70%, where efficacy is the standard measure: $(1 - R_v/R_u) \times 100\%$, with R_v and R_u the attack rates among the vaccinated and unvaccinated, respectively. Second, we had to account for the intra-unit correlation. We took a quasi-likelihood model for the basic approach, in which a marginal Poisson regression model is fit, and then the estimated covariance matrix of the parameters is multiplied by an estimate of the overdispersion parameter σ^2, usually estimated from the Pearson χ^2 residuals [23]. This parameter represents variability beyond that expected under an assumption of a Poisson distribution with the same rate parameter across randomization units that are in the same treatment group. To obtain an initial estimate of σ^2, we took pneumococcal surveillance data from 1988–1995 on the Navajo Nation and fit a Poisson regression model accounting for year and provider (the eight IHS service units and one private hospital). The estimate of σ^2 was 1.85. In the usual moment-based model of overdispersion, the design effect (or variance inflation factor) is $\sigma^2 = 1 + (n - 1)\rho$, where n is the number of participants in a given randomization unit (given the units are the same size)

randomization, and ρ is the intra-unit correlation. Since we actually have 36 randomization units on the Navajo Nation instead of the nine that 1.85 was based upon, we set $(n - 1)\rho = \sigma^2 - 1 = 0.85$ and multiplied by 9/36, obtaining 0.21, for a final estimate of 1.21 for σ^2. This calculation assumes ρ to be the same regardless of whether one is considering the provider (approximately service unit) level or the final randomization unit level (about a fourth the population size of the provider level). In general one might expect ρ to be greater when smaller units are employed, even though σ^2 would be decreasing with decreasing unit size, reaching unity in an individually randomized trial. We did not anticipate much difference in ρ when going from the provider level to the randomization unit level because of the criteria we used to create the randomization units. The goal was to not make the units so small that significant amounts of social interaction would be taking place across the units. Socioeconomic status is not highly variable on the reservations, but it could be that differing rural-to-urban residence ratios between different size units could increase ρ when going toward smaller units. As a safeguard against this possibly anticonservative assumption regarding the sample size, the plan was to take a mid-study look at the design effect σ^2 as stated below. We had some leeway to increase the length of the study to accrue more cases if necessary.

Thus, for any sample size that would be required for an individually randomized trial, we just multiply by 1.21 to get the requisite sample size for our group-randomized trial. The resultant sample size was thus $40(1.21) = 48$ observed cases (the original 40 was based on a slightly conservative exact calculation, thus we took the liberty of rounding down).

To estimate the amount of time that would be needed to obtain these 48 cases, we took historic incidence data and current estimates of birth cohorts and marched cohorts of infants through time via a Leslie matrix [24], exposed to half-yearly age-specific incidence rates. We used enrollment, loss to follow-up, and incidence data from the period of a Hib vaccine efficacy trial carried out in the Navajo population from 1988–1991. The background incidence rates of invasive pneumoccocal disease were 808, 757, and 470 per 100,000 child-years in the intervals 6–11, 12–17, and 18–23 months of age; loss to follow-up was estimated globally to be 20% of child-years, although it would have been more accurate to use age-specific loss rates. With these values, we estimated 28 months would be necessary to complete the enrollment and surveillance period of the trial to achieve the goal of 48 primary efficacy cases, during which time about 9000 infants would be enrolled in the primary efficacy cohort.

Interim Analysis

We proposed performing one interim analysis using a Lan-DeMets implementation of the O'Brien-Fleming boundary rule [25]. In the absence of overdispersion, we would take the first look after 17 primary efficacy cases have accrued. To cross the boundary requires that the *p*-value relative to a null hypothesis of 0% efficacy be less than 0.00239 (two-sided), leaving a size 0.049 test at 40 cases. This boundary would be crossed with a 15:2 or more extreme case split, which we would have 80% power to achieve if in fact the true efficacy were 90%. Note that our overall sample size objective was to have a lower bound of not 0%, but 20%. Yet if we had a 15:2 split in an individually random-

ized trial, the exact 95% confidence interval would be (42.7%, 98.5%), which we thought would be sufficient. As an approximation to account for the penalty incurred by the above interim look, a 95.1% interval could be calculated. Because of overdispersion, we would actually look at the data after accruing 1.21 · 17, or 21, cases. If the first 17 cases were all in distinct randomization units, however, there would be no overdispersion and we could perform the interim analysis at that point in time.

We planned to modify the above approach to guard against a situation of greater-than-expected overdispersion. When 21 cases have accrued, we would have an independent statistician run a Poisson regression with only a treatment covariate and report only the estimate of σ^2. If the estimate were less than or equal to 1.21, we would conduct the interim analysis; if it were greater, the requisite sample size for the interim analysis would be recalculated. No information on treatment effect is conveyed, since under this model the mean and overdispersion are independent.

Analytic Approach

The basic model we have assumed both for planning and analysis of the trial is an overdispersed Poisson regression model. It is a quasi-likelihood approach in which a Poisson regression model is fit at the randomization unit level with a scale factor adjustment for overdispersion. For the case of the single treatment covariate for vaccine effect, this can be fit in a simple two-step process. In the first pass through the data, we accumulate the total numbers of events in the unvaccinated ($j = 0$) and vaccinated ($j = 1$) groups in each of the $i = 1, \ldots, m$ randomization units within each group (for a total of $2m$ randomization units):

$$y_{\bullet j} = \sum_{i}^{m} y_{ij}$$

and the total person-time of follow-up in each group:

$$t_{\bullet j} = \sum_{i}^{m} t_{ij}.$$

In the second pass, we get: $\hat{y}_{ij} = \exp(\hat{\alpha} + j\hat{\beta} + \ln t_{ij})$,

$$s_{\hat{\beta}} = \sqrt{y_{\bullet 0}^{-1} + y_{\bullet 1}^{-1}},$$

and the estimated overdispersion

$$\hat{\sigma}^2 = \frac{1}{2m-2} \sum_{i,j} (y_{ij} - \hat{y}_{ij})^2 / \hat{y}_{ij},$$

where $\hat{\alpha} = \ln(y_{\bullet 0}/t_{\bullet 0})$ and $\hat{\beta} = \ln(y_{\bullet 1}/t_{\bullet 1}) - \ln(y_{\bullet 0}/t_{\bullet 0})$. Then vaccine efficacy is calculated as: $[1 - \exp(\hat{\beta})] \times 100\%$ with 95% CI: $[1 - \exp(\hat{\beta} \pm 1.96 s_{\beta}\hat{\sigma})] \times 100\%$. Note that a disadvantage of this approach is that if a $y_{\bullet j}$ is zero (i.e., there are no cases in one of the trial arms), the variance cannot be estimated. If this were to happen, two alternative strategies would include: (1) calculate a permutation test-based confidence interval; and (2) fit a negative binomial regression model and either graph the profile likelihood function for the treatment parameter or

calculate a likelihood ratio test (the treatment parameter will have an infinite estimate, precluding a Wald-based interval or test).

DISCUSSION

In this article we have focused on the novel aspects of this trial relating to its group-randomized design. Such designs require close attention to the formation of the treatment randomization units. In some trials, the units may be obvious or given, for example, geographically separated cities as in the COMMIT smoking cessation trial [26]. In others, several factors need to be balanced: logistical ease of determining unit boundaries and administering treatments, formation of enough units to achieve adequate statistical power, and keeping units from being so numerous in a circumscribed geographic area as to promote cross-unit contamination. The first and third of these factors were major factors in the decision of Jaffar et al. [27] not to undertake a group-randomized trial in the Gambia; the last led Wawer et al. [28] to group their numbers of communities from 56 into ten randomization units. None of the other five principal investigators who have been planning pneumococcal vaccine trials have considered randomization by group to be a feasible alternative in their study populations. Nevertheless, the relatively contained nature of the Native American populations in our study, combined with the developed country advantages of easily determined identity and residence, were among the factors that enabled us to conduct such a trial. After our trial began, a large group-randomized study of Hib conjugate vaccine was begun in Lombok, Indonesia, although the primary goal of that study is not vaccine effectiveness per se, but to estimate the burden of Hib-associated disease.

The analysis of group-randomized trials, although it has accumulated a substantial literature, still has some problems outstanding. In our trial design, anticipated high efficacy (as opposed to moderate effects of behavioral community interventions) and effectively small sample size can render problematic the use of methods based on large-sample theory. In such situations, permutation analyses may be useful, but their properties under an overdispersed Poisson model with differential follow-up have not been investigated.

The principal outcome of this study, the measure of the total effect of the vaccine when deployed in a community, is conditional on the uptake on the part of the study population. A licensed, recommended vaccine would most likely lead to higher coverage rates in communities with potentially greater indirect and hence total effects. From this perspective, we might consider our estimate as a lower bound on the advantage of routine use of the vaccine, both in terms of total and overall reduction of invasive pneumococcal disease due to the serotypes included in the vaccine. Alternatively, if the main transmission pattern is from much older siblings and adults to children under 2 years of age, then it may take many years for the full impact of reduced carriage to be manifested, which our trial could not hope to estimate. Still, for vaccines that induce some degree of indirect effects, these estimates will give a better idea of the impact of a vaccine than the direct effects that are estimated in individually randomized vaccine efficacy trials.

ADDENDUM

On February 17, 2000, the FDA approved licensure of the 7VPnC vaccine, largely based on the results of the northern California trial [13]. At that time, our trial had accrued nine cases in the primary efficacy analysis group. The Data and Safety Monitoring Board for the trial, after receiving and analyzing further data, recommended in June 2000 that steps be taken to immunize control group participants (MnCC recipients) with a commercial lot of 7VPnC. The study code was broken in October 2000, at which time control study participants began receiving the 7VPnC series. The IHS began routine immunization of infants and children on the Navajo and White Mountain Apache Tribe reservations with 7VPnC in late October and November of 2000. Surveillance for pneumococcal disease is being maintained, and all data collected through the end of 2001 will be analyzed for evidence of indirect effects. At that time, the last infants enrolled in the study will have reached 2 years of age. Additional information on uptake and coverage with the licensed vaccine among nonstudy children will be collected to carry out this analysis.

This work was supported in part by grants from Wyeth Lederle Vaccines, USAID (HRN-A-00-96-9006), WHO, and CDC. The opinions of the authors are not necessarily those of the Indian Health Service. We thank the Health Boards of the White Mountain Apache Tribe; the Health Boards of Chinle (Arizona), Fort Defiance (Arizona), Gallup (New Mexico), and Shiprock (New Mexico) of the Navajo Nation; the Navajo Nation Health Research Review Board; the Phoenix Area Indian Health Service Institutional Review Board; the National Indian Health Service Institutional Review Board; and the Johns Hopkins University Joint Committee on Clinical Investigation for review and approval of the study. We are indebted to Larry Rodgers for his invaluable assistance in defining the randomization units and producing maps.

REFERENCES

1. Istre GR, Tarpay M, Anderson M, Pryor A, Welch D. Invasive disease due to *Streptococcus pneumoniae* in an area with a high rate of relative penicillin resistance. *J Infect Dis* 1987;156:732–735.
2. Breiman RF, Spika JS, Navarro VJ, Darden PM, Darby CP. Pneumococcal bacteremia in Charleston County, South Carolina. A decade later. *Arch Intern Med* 1990; 150:1401–1405.
3. Greenwood B. The epidemiology of pneumococcal infection in children in the developing world. *Philos Trans R Soc Lond B Biol Sci* 1999;354:777–785.
4. The WHO Young Infants Study Group. Bacterial etiology of serious infections in young infants in developing countries: Results of a multicenter study. *Pediatr Infect Dis J* 1999;18:S17–S22.
5. Garenne M, Ronsmans C, Campbell H. The magnitude of mortality from acute respiratory infections in children under 5 years in developing countries. *World Health Stat Q* 1992;45:180–191.
6. Institute of Medicine. New vaccine development: Establishing priorities, Volume II. *Diseases of Importance in Developing Countries.* Washington DC: National Academy Press; 1986:357–375.
7. Advisory Committee on Immunization Practices. Prevention of pneumococcal disease: Recommendations of the Advisory Committee on Immunization Practices (ACIP). *MMWR Morb Mortal Wkly Rep* 1997;46(RR-08):1–24.
8. Duchin JS, Breiman RF, Diamond A, et al. High prevalence of multidrug-resistant Streptococcus pneumoniae among children in a rural Kentucky community. *Pediatr Infect Dis J* 1995;14:745–750.

9. Davidson M, Schraer CD, Parkinson AJ, et al. Invasive pneumococcal disease in an Alaska Native population, 1980 through 1986. *JAMA* 1989;261:715–718.
10. Cortese MM, Wolff M, Almeido-Hill J, et al. High incidence rates of invasive pneumococcal disease in the White Mountain Apache population. *Arch Intern Med* 1992;152:2277–2282.
11. Davidson M, Parkinson AJ, Bulkow LR, et al. The epidemiology of invasive pneumococcal disease in Alaska, 1986–1990: Ethnic differences and opportunities for prevention. *J Infect Dis* 1994;170:368–376.
12. O'Brien KL, Croll J, Parkinson AJ, Reid R, Santosham M. Active laboratory based surveillance for invasive *Streptococcus pneumoniae* (pneumococcus) among Navajo people in the American Southwest, 1989–1996 (abstract). The 39th Interscience Conference on Antimicrobial Agents and Chemotherapy, San Francisco, 1999.
13. Black S, Shinefield H, Fireman B, et al. Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. Northern California Kaiser Permanente Vaccine Study Center Group. *Pediatr Infect Dis J* 2000;19:187–195.
14. Hayes RJ, Alexander NDE, Bennett S, Cousens SN. Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Stat Methods Med Res* 2000;9:95–116.
15. Galil K, Singleton R, Levine OS, et al. Reemergence of invasive *haemophilus influenzae* type b disease in a well-vaccinated population in remote Alaska. *J Infect Dis* 1999;179:101–106.
16. van Alphen L, Spanjaard L, van der Ende A, Schuurman I, Dankert J. Effect of nationwide vaccination of 3-month-old infants in The Netherlands with conjugate Haemophilus influenzae type b vaccine: High efficacy and lack of herd immunity. *J Pediatr* 1997;131:869–873.
17. Moulton LH, Chung S, Croll J, et al. Estimation of the indirect effect of Hib conjugate vaccine in an American Indian population. *Int J Epidemiol* 2000;29:753–756.
18. Halloran ME, Haber M, Longini IM, Jr., Struchiner CJ. Direct and indirect effects in vaccine efficacy and effectiveness. *Am J Epidemiol* 1991;133:323–331.
19. Halloran ME, Struchiner CJ, Longini IM, Jr. Study designs for evaluating different efficacy and effectiveness aspects of vaccines. *Am J Epidemiol* 1997;146:789–803.
20. Fine PE. Herd immunity: History, theory, practice. *Epidemiol Rev* 1993;15:265–302.
21. Peduzzi P, Donta S, Iwane M. A review of the design of vaccine efficacy trials and a proposal for the design of the VA Cooperative Study of Active Immunotherapy of HIV Infection. *Control Clin Trials* 1997;18:397–419.
22. Bronsdon MA, O'Brien KL, Facklam RR, Schwartz B. Immunoblot detection of multiple serotypes of Streptococcus pneumoniae (Pnc) in nassopharyngeal (NP) specimens using monoclonal antibodies (MAb) (abstract). The Second International Symposium on Pneumococci and Pneumococcal Disease, Sun City, South Africa, Abstract P96, 2000.
23. McCullagh P, Nelder JA. *Generalized Linear Models*. 2nd ed. New York: Chapman and Hall; 1989.
24. Leslie PH. Some further notes on the use of matrices in population mathematics. *Biometrika* 1948;35:213–245.
25. Cytel Software Corporation. EaSt (2.0). Cambridge, Massachusetts: Cytel Software Corporation; 1992.
26. COMMIT Research Group. Community Intervention Trial for Smoking Cessation (COMMIT): Summary of design and intervention. *J Natl Cancer Inst* 1991;83:1620–1628.
27. Jaffar S, Leach A, Hall AJ, et al. Preparation for a pneumococcal vaccine trial in The Gambia: Individual or community randomisation? *Vaccine* 1999;18:633–640.
28. Wawer MJ, Gray RH, Sewankambo NK, et al. A randomized, community trial of intensive sexually transmitted disease control for AIDS prevention, Rakai, Uganda. *AIDS* 1998;12:1211–1225.