Combination Vaccines: Issues in Evaluation of Effectiveness and Safety

Harry A. Guess

INTRODUCTION

Combination vaccines, which help provide protection against two or more diseases or against multiple serotypes of a single disease are increasingly used to help reduce the number of injections required for childhood immunization (1). Familiar combinations include measles, mumps, and rubella (MMR); diphtheria, tetanus, and whole-cell pertussis (DTP); poliomyelitis (serotypes 1–3); pneumococcal disease (23 serotypes); and influenza A and B. Recently a number of diphtheria, tetanus, and acellular pertussis (DTaP) vaccines have received licensing approval, and DTaP is now recommended as the preferred combination in the United States, with DTP vaccine as an acceptable alternative (2). In the United States there have been only two recently licensed vaccines combining moieties that had previously required separate injections. A combination vaccine for \textit{Haemophilus influenzae} type b (Hib) and hepatitis B (COMVAX®, Merck & Co., Inc., Whitehouse Station, NJ) is available for use beginning at 2 months of age and a combination DTaP/Hib vaccine (TriHIBit®, Pasteur Mérieux Connaught, USA, Swiftwater, PA) is available for use at 15 months of age. Further combinations involving DTaP, Hib, inactivated poliomyelitis vaccine (IPV), and hepatitis B are licensed in Canada and European countries, and additional combinations, including MMR and varicella, are in clinical trials.

Regulatory guidance on formulation, manufacturing, preclinical testing, and pre-licensure clinical trials of combination vaccines has recently been provided by the US Food and Drug Administration (FDA) (3).

While the immunologic, manufacturing, and regulatory aspects of combination vaccines have been the subject of a number of reviews, less has been written on the evaluation of effectiveness and safety. In what follows, we will review the recent literature on the epidemiologic evaluation of effectiveness, safety, and potential improvements in vaccine acceptance and coverage associated with combination vaccines.

SEROLOGIC RESPONSES AND THEIR RELATION TO CLINICAL EFFICACY OF COMBINATION VACCINES

There are a number of physical, chemical, and immunologic mechanisms by which serologic responses to antigens in combination vaccines may differ from those obtained with separate administration of the components (3, 4). Preservatives used with one component may alter the potency of other components. For example, thimerosal, a preservative in DTP vaccines, has long been known to adversely affect the potency of IPV's (5, 6). Buffers used with different components may be incompatible (4). Antibody titers to some live viruses may be lower when administered in a combination vaccine than when administered separately (7). Reduced antibody responses have also been shown when multiple protein-conjugated vaccines sharing common protein epitopes have been administered simultaneously (8). Some combination vaccines and vaccine mixtures containing Hib antigens have been shown to have a lower Hib antibody response than has been seen with separate administration, although the clinical implications of the lower responses have not been established (9–12).

Serologic correlates of efficacy have been established for a number of diseases (3, 13). Several lines of indirect evidence indicate that serum neutralizing antibodies to poliovirus induced by natural infection, oral poliomyelitis vaccine (OPV) or IPV are protective against paralytic disease (5). Antibodies to the hepatitis B surface antigen induced by vaccination or by natural disease have been correlated with protection on the basis of animal challenge experiments and vaccine clinical trials in humans (14, 15). Serum bacteriocidal activity against \textit{Neisseria meningitidis}, serum anti-
toxin activity against diphtheria and tetanus, antibody to the hepatitis B surface antigen, and antibodies to serotype-specific capsular polysaccharides of Streptococcus pneumoniae have all been shown to be correlated with protection (13-16). Antibodies to the Hib capsular polysaccharide (polyribosyl ribitol phosphate (PRP)) have also been correlated with protection (17).

However, as noted by the FDA (3) and by Granoff and Lucas (17), the quality and not simply the quantity of the antibody also needs to be considered in comparative immunogenicity studies of combination vaccines. For example, antibodies with a low functional affinity for the Hib-PRP show lower percent binding at low antigen concentrations (a measure of low avidity). This suggests that the extent to which a given level of vaccine-induced antibody to PRP protects against clinical Hib disease may be influenced by the avidity of the antibody (17). In addition, recent experience with acellular pertussis vaccines shown to be effective in protecting against clinical disease has indicated that levels of antibody induced by vaccines are not predictive of protection (13, 18). Several markers of cell-mediated immunity to pertussis antigens may ultimately prove more useful than serology in predicting vaccine-induced protection against pertussis (13, 19).

Thus, factors other than antibody levels or seroconversion rates (percentage of vaccinees achieving a previously established “protective” level of antibody) need to be considered in the immunologic evaluation of combination vaccines (3). For example, markers of immunologic memory have proven to be important for evaluating the potential for long-term protection of some vaccines (20-22). One way in which evidence of immunologic memory can be provided is through demonstrating production of an anamnestic immunoglobulin G antibody response to revaccination or exposure to natural disease (23). Antibody avidity assays have also been used as surrogate markers of immunologic memory (24). Hence, even if there are well-established protective antibody levels applicable to the individual antigens comprising a new combination vaccine, it would be important to provide evidence that the previously established protective levels can be validly applied to the new combination vaccine.

PREAPPROVAL SAFETY AND EFFICACY EVALUATION OF COMBINATION VACCINES

Formulation and testing requirements

Prior to any human clinical evaluation of new components of combination vaccines, extensive preclinical testing is required to ensure compatibility and stability of the proposed new formulations, as well as consistency of the final manufacturing process (25). When there are suitable animal models for the diseases, both the individual components and their combination should be tested for animal immunogenicity and for protective efficacy in animal challenge studies (3). Safety testing in animals will also generally be required for any component for which such data are not already available. If adjuvants are used that have not been used previously in currently licensed vaccines, animal toxicity testing of the adjuvants may be required (3). The preclinical development of any combination vaccine is typically carried out as an iterative process, as formulations are modified and refined on the basis of in vitro and in vivo assays.

Preapproval testing of combination vaccines for tolerability and immunogenicity in man

Each new component will generally be initially evaluated individually for immunogenicity and tolerability in adults, unless the new component has already undergone human clinical testing in the formulation to be used in the new combined vaccine. Once this initial testing has shown a new component to be immunogenic and well tolerated, it will then typically be evaluated in a clinical immunogenicity trial where it is administered separately but simultaneously with the existing vaccine. All of this should be done before the new component is tested as part of a combined vaccine. It is important to recognize that from a regulatory perspective, a combination vaccine is a new vaccine, which must undergo regulatory review and licensing approval, even if all of the moieties are from approved single-component vaccines (3).

On purely statistical grounds one could argue that 2N arms would be required in a clinical trial to evaluate immunogenicity of an N-component combination vaccine if there was no biologic basis for eliminating any of the theoretically possible interactions from consideration (26). In practice, it is rarely necessary to test all possible combinations of components. Most comparative clinical immunogenicity trials of combined vaccines are conducted with previously approved combinations and/or separate components as active controls. Nonetheless, the statistical multiplicity in comparing immune responses to large numbers of different components can lead to very large sample sizes and can sometimes make apparent differences in immunogenicity of individual components difficult to interpret.

Evolving standards for the conduct of equivalence trials provide a paradigm for vaccine studies to establish serologic equivalence between a given combination vaccine and the individual component antigens of which it is composed (27, 28). A common problem is
to establish that the seroconversion rate for a vaccine component administered as a combination ($R_c$) is not lower by a predefined clinically important amount, $\Delta$, than that obtained when the component is administered separately ($R_s$). Typically, the magnitude of the amount, $\Delta$, will have been agreed upon in discussions with the FDA before the trial is undertaken.

As outlined by Blackwelder (27), the statistical test to establish non-inferiority may be formulated by taking the null hypothesis to be $H_0: R_c \geq R_s - \Delta$. Rejection of $H_0$ at a significance level of $\alpha$ is required to conclude that $R_s > R_c - \Delta$. Hypothesis tests regarding comparisons of geometric mean titers (GMTs) would be formulated similarly, except that it may be more appropriate to test for equivalence rather than non-inferiority, with an obvious modification of the above equations. If it were necessary to compare other immunologic parameters such as avidity or markers of cell-mediated immunity, the same type of statistical considerations would apply. Because all of these comparisons would typically need to be established simultaneously for each component of a combination vaccine, the required sample sizes can be quite substantial.

Throughout the entire process of preapproval clinical trials, there will be extensive monitoring for local reactions, systemic complaints, and serious adverse events. Even if all of the components of the combination are themselves licensed vaccines, it is still necessary to evaluate the safety and tolerability profiles of the combination.

Evidence of clinical efficacy required for approval of combination vaccines

If immunologic correlates of efficacy are well established for each component of a combination vaccine, data from immunogenicity trials could provide a basis for license approval without the need for additional evidence studies (3). When immunologic data are not sufficient, the type of evidence needed for approval will depend on the specific combination vaccine under consideration. For example, recent approvals of combination vaccines containing diphtheria, tetanus, and acellular pertussis antigens required randomized, double-blind, clinical efficacy trials to demonstrate that the new pertussis components were protective against clinical disease. However, immunologic data were relied upon to show that efficacy of the diphtheria and tetanus antigens had been maintained.

Clinical efficacy trials of combination vaccines for multiple serotypes may have as a primary endpoint the aggregate of disease with all serotypes included. However, the study should be sufficiently powered to permit meaningful subgroup analysis of protection against some individual serotypes (3). It is important to consider differences between the serotype distribution in the population in which efficacy trials are conducted and those in which the vaccine will subsequently be used. Such differences can lead to important differences between the efficacy shown in clinical trials and that found with use of the vaccine in clinical practice after licensure.

**EPIEMIOLOGIC STUDIES OF COMBINATION VACCINE EFFECTIVENESS IN CLINICAL PRACTICE**

**Efficacy and effectiveness**

Once a vaccine has been licensed on the basis of preclinical studies and clinical trials, demonstrating its safety and efficacy, it is important to evaluate its effectiveness in clinical practice (29, 30). Efficacy of a vaccine refers to the reduction in disease measured in a carefully-monitored, randomized controlled clinical trial conducted in a homogeneous population according to a defined protocol. Effectiveness refers to the reduction in disease measured under conditions of use of the vaccine in ordinary clinical practice (31). For most pharmaceutical interventions, effectiveness would be expected to be somewhat less than efficacy. The original polysaccharide Hib vaccines may have been an example of a type of vaccine whose effectiveness was less than its efficacy.

On the other hand, Hib conjugate vaccines represent at least one type of intervention whose field effectiveness in elimination of invasive Hib disease may actually be greater than what would have been predicted from immunization rates and efficacy in clinical trials (32). It appears that widespread vaccine use decreased the transmission of the disease to non-vaccinees, thereby causing a decreased likelihood of disease among both vaccinees and non-vaccinees. A possible mechanism is through lower nasopharyngeal carriage of the organism in vaccinated children than in unvaccinated children (33). This example of a vaccine whose effectiveness may be greater than its efficacy provides a further illustration of the importance of monitoring vaccine effectiveness after licensure, as emphasized by Orenstein et al. (34).

**Commonly used designs for observational studies of vaccine effectiveness**

Epidemiologic studies of vaccine effectiveness in clinical practice generally involve either direct (cohort) or indirect (case-control) comparison of the incidence of disease in vaccinated and non-vaccinated individuals using standard methods that have been well described in a number of reviews (29, 34). A
design more commonly used in health policy research than in epidemiology is the “interrupted time series” (35). This has been used to estimate the effect of policy changes on health outcomes, while controlling for temporal trends (36). A similar method has recently been used to estimate the effects of interrupted pertussis vaccination programs on disease rates by comparing incidences in countries which have maintained high pertussis immunization rates with incidences in countries in which program interruptions have occurred (37).

Variation in effectiveness with time since vaccination or season of vaccination has been studied not only with traditional cohort designs but also with case-control designs (38, 39) and, more recently, with a nonparametric design (40). More complex mathematical modeling has been used to evaluate potential long-term epidemiologic effects of new immunization policies (41, 42) and to predict disease outbreaks in partially immunized populations (43).

Because the efficacy of most vaccines is typically at least 50 percent, effects of selection bias, unmeasured confounding, misclassification, and other non-random sources of error are typically capable of being overcome by careful study design. For example, a randomized, double-blind, placebo-controlled clinical trial of influenza vaccination in healthy elderly patients in Holland (44) and a retrospective cohort study of influenza vaccine effectiveness among residents of a Minnesota health maintenance organization (HMO) (45) produced very similar measures of overall vaccine effectiveness. The randomized trial used a more rigorous study design and a more stringent case definition. However, the observational study was able to evaluate not only more clinically important endpoints, including hospitalization and mortality, but also was capable of quantifying the cost-effectiveness of the intervention. These two studies, which were undertaken quite independently, illustrate how observational studies and randomized clinical trials can be used together to provide additional information about vaccine effectiveness.

Potential problems to be faced in comparative observational effectiveness studies

Observational studies to quantify differences in effectiveness of two vaccines are typically much more sensitive to bias than observational studies of the absolute effectiveness of a vaccine, because the effect size is typically larger in the latter type of study. Consequently, it is difficult to use observational studies to determine whether differences in surrogate markers of efficacy between two licensed combination vaccines actually translate into differences in effectiveness. For example, study design differences have made it difficult to interpret conclusions from observational effectiveness studies of several acellular pertussis vaccines (46). As was discussed in the previous section, it is difficult enough to conduct an equivalence trial of a multicomponent vaccine using a randomized design, where the main problem lies in the large sample sizes required. Observational designs are typically only useful when comparing vaccines with very large differences in effectiveness. For example, it may be possible to use observational studies to demonstrate an advantage of one vaccine over another in providing greater protection after the first dose of a series. Large simple randomized clinical trials continue to be by far the preferred approach for providing valid estimates of moderate differences in effectiveness between two efficacious vaccines.

EPIDEMIOLOGIC RESOURCES FOR SAFETY STUDIES OF COMBINATION VACCINES

The National Childhood Vaccine Injury Act of 1986 requires health care providers who administer vaccines in the United States to maintain permanent vaccination records and to report occurrences of certain adverse events to the Vaccine Adverse Events Reporting System (VAERS) (47). The annual number of reports now exceeds 10,000 per year (48). While this system of passive reporting provides a basis for identifying potential vaccine-related adverse events, it is not well suited to analytical studies. Despite the requirement for reporting, there is a substantial degree of under-reporting in VAERS compared with reporting in controlled studies (49).

Rates of underreporting in VAERS also vary according to a number of unknown factors, creating the potential for bias in comparisons of vaccine safety based on passive reporting in VAERS. To evaluate this phenomenon, a retrospective cohort study was performed to compare rates of adverse events in children who had received one of two recombinant hepatitis B vaccines (50).

In VAERS, rates of serious adverse events were approximately three times higher with one vaccine than the other ($p < 0.01$). A serious adverse event was defined to be one that resulted in hospitalization, permanent disability, or death, or was judged to be life threatening. The cohort study was conducted with the Centers for Disease Control and Protection (CDC) Vaccine Safety Datalink Project, which makes use of linked medical records and vaccine records on more than 500,000 children aged 0–6 years enrolled at four large HMOs in the western United States (48). In contrast to the pronounced difference in reporting rates seen in VAERS, the rates of all hospitalizations within 30 days after receiving either of the two vaccines were
nearly equal (relative risk = 1.04, 95 percent confidence interval: 0.93, 1.15). An additional analysis addressed hospitalizations or emergency room visits for diagnostic categories that represented either events commonly reported to VAERS after hepatitis B vaccination in infants or else preselected by the investigators as being possibly vaccine related. In this analysis the rates were again essentially equal (relative risk = 1.07, 95 percent confidence interval: 0.94, 1.22). This study underscores the importance of undertaking analytic epidemiologic studies to evaluate findings from passive surveillance systems.

The Vaccine Safety Datalink Project was funded by CDC in 1991 and represents the largest comprehensive database resource for analytic studies of vaccine safety in the United States (48). Initially the project focused on children aged 0–6 years but it has been expanded to include adolescents and adults. Health service use information for each subject is computerized on an individual patient basis and organized into files containing demographics and enrollment, vaccine administration records (including date, manufacturer, lot number, vaccination site, and whether the vaccinations were obtained within the HMO or from outside providers), hospital and emergency room visits (at all four sites) and outpatient visits (at two sites), selected diagnostic procedures and laboratory data, prescription drugs, linkage to state birth and death certificates, geocoding (for estimation of socioeconomic status based on census block codes), and past medical diagnoses. Written medical records are available for review. Numerous data quality control procedures have been instituted to ensure accuracy and completeness of the information and to monitor and maintain quality of the records abstraction process. Software has been developed to permit computing person-time at risk in various defined time windows after vaccination. A number of specific analytic methodologies have been employed, depending on the specific question to be addressed.

Safety studies of combination vaccines published to date have included 1) the risk of chronic arthropathy among women following rubella vaccination (51), 2) rates of serious clinical events in recipients of MMR vaccine at 4–5 years and 10–12 years (52), 3) timing of seizures following DTP vaccination and MMR vaccination (43), and 4) the risk of hospitalization because of aseptic meningitis after MMR vaccination (53).

The VAERS and the CDC Vaccine Safety Datalink Project complement each other. VAERS provides a mechanism for early identification of potential safety problems with new vaccines used singly or in combination with existing vaccines. The Safety Datalink Project provides a mechanism for analytic studies, quantitative risk estimation, and comparative studies. Together they represent a unique national resource for monitoring vaccine safety. In particular, these resources are especially useful for studying the safety of combination vaccines and simultaneous use of multiple vaccines as they are administered in clinical practice.

SIMULTANOUS ADMINISTRATION OF VACCINES

Preapproval studies of new vaccines commonly include data confirming safety and immunogenicity when the new vaccine is administered simultaneously with licensed vaccines that would be given on a similar or overlapping schedule (3). However, it is impractical to conduct preapproval studies of all combinations that are used in clinical practice. The American Academy of Pediatrics Committee on Infectious Disease states that, "Because simultaneous administration of common vaccines is not known to affect the efficacy or safety of any of the routinely recommended childhood vaccines, if return of a vaccine recipient for further immunization is doubtful, simultaneous administration of all vaccines (DTaP [or DTP], OPV or IPV, MMR, rubella, varicella, hepatitis B, and Hib vaccines) appropriate for age and previous vaccination status of the recipient is recommended. Simultaneous administration of multiple vaccines can raise immunization rates significantly" (54, p. 21). Simultaneous administration of all vaccines for which a child is eligible at the time of the visit is also one of the Standards for Pediatric Immunization Practices recommended by the National Vaccine Advisory Committee and approved by the US Public Health Service (55). Ecologic analyses of reports from passive surveillance systems for monitoring reports of adverse events among vaccinees suggest that simultaneous administration of common pediatric vaccines do not appear to have been associated with increased reports of serious adverse events (56).

The comment that simultaneous administration of multiple vaccines can raise immunization rates is supported by a decision analysis performed in the Northern California Kaiser Permanente Medical Group (57). Analysis of charts of 4,691 children who had missed one or more immunizations due in their second year showed that about one third of underimmunized children in the Northern California Kaiser Permanente Medical Group would have received all their second year immunizations if their providers had followed the guideline to administer simultaneously all vaccine doses for which the child was eligible. Despite this finding and the above recommendations, a national survey of a random sample of pediatricians and family practitioners found that approximately one third would not vaccinate an 18-month-old healthy

Epidemiol Rev Vol. 21, No. 1, 1999
child with all four vaccines because of perceived medical contraindications, concerns about pain, parental objections, or costs (58). It appears likely that increased use of combination vaccines can help overcome some of the reluctance of practitioners to administer multiple vaccines simultaneously by reducing the number of injections necessary to deliver childhood vaccinations.

CONCLUSIONS

Combination vaccines are expected to become increasingly common and increasingly complex over the next decade as additional antigens are added to the pediatric vaccination regimen (59, 60). This should permit a further reduction in childhood infectious diseases and lessen the potential dangers of increasing antibiotic resistance among organisms causing serious childhood disease. At the same time it will be necessary to develop methodology to know how to use the vaccines most efficiently and to monitor their safety, effectiveness, and effects on immunization coverage in practice.

REFERENCES

52. Davis RL, Marcuse E, Black S, et al. MMR2 immunization at 4 to 5 years and 10 to 12 years of age: a comparison of adverse clinical events after immunization in the Vaccine Safety Datalink project. The Vaccine Safety Datalink Team. Pediatrics 1997;100:767-71. [See comments in Pediatrics 1997;100:891-2.]