Evaluating the Safety of New Vaccines: Summary of a Workshop

Susan S. Ellenberg, PhD, Mary A. Foulkes, PhD, Karen Midthun, MD, and Karen L. Goldenthal, MD

Public concerns about the safety of vaccines arise on a regular basis. In November 2000, a workshop titled "Evaluation of New Vaccines: How Much Safety Data?" was convened by US Public Health Service agencies, including the Food and Drug Administration, the National Institutes of Health, the Centers for Disease Control and Prevention, and the Health Resources and Services Administration, to discuss appropriate methods for evaluating the safety of new vaccines.

Workshop presentations addressed the current standards and approaches for new vaccine evaluation and postlicensure surveillance, as well as public views about vaccine safety and alternative approaches that could be considered.

The advantages and disadvantages of conducting large controlled trials before licensure or widespread use of a new vaccine were discussed. We summarize these presentations and discussions. (*Am J Public Health.* 2005;95:800–807. doi:10.2105/AJPH.2004.039438)

Public concerns about the safety of vaccines arise regularly. Available data suggest that such concerns are seldom warranted; however, the amount of data collected before licensure of a vaccine—particularly in randomized controlled trials—varies greatly and is usually too limited to allow an assessment of the vaccine's potential association with rare but serious adverse events. In 1999, for example, a new vaccine designed to protect infants against rotaviral disease was withdrawn from the market by its manufacturer less than a year after licensure because of an association with intussusception, a potentially life-threatening condition.[1,2]

In November 2000, a workshop was convened by agencies of the US Public Health Service, including the Food and Drug Administration (FDA), the National Institutes of Health, the Centers for Disease Control and Prevention (CDC), and the Health Resources and Services Administration, to discuss appropriate methods for evaluating the safety of new vaccines. Workshop presentations addressed the current standards and approaches for new vaccine evaluation and postlicensure surveillance, public views about vaccine safety, and alternative approaches that could be considered. In particular, discussion focused on the advantages and disadvantages of conducting larger controlled trials to obtain information on serious but less common adverse events prior to licensure or widespread use of a new vaccine. Specific issues considered included the following:

- What is the correct balance between the amount of data collected on efficacy and on safety?
- What amount and types of safety information should be available before licensure or recommendation of a vaccine for widespread use?
- What size difference in the rate of a serious adverse event would be large enough to be clinically significant and should therefore be detectable in a controlled study?

Invited speakers and discussants included representatives of government agencies and vaccine manufacturers; academic scientists with expertise in infectious disease, vaccine development, and bioethics; and members of public advocacy groups. The workshop agenda included several periods of extended audience discussion. Here we summarize these presentations and discussions. To reflect the proceedings accurately, we limit attention to issues that were raised and discussed at the workshop.

These issues continue to be highly relevant to vaccine developers and the public health community. Public concerns about vaccine safety have not diminished. Since the time of the workshop, however, new concerns about susceptibility to infectious diseases have arisen. The possibility of an influenza pandemic or bioterror attacks, as well as the emergence of new potentially lethal infectious diseases such as severe acute respiratory syndrome (SARS), makes it clear that vaccine development will never follow a "one size fits all" pattern; balancing the benefits against the risks of introducing new vaccines will remain very much a case-by-case consideration.

SESSION 1: CURRENT APPROACHES AND PUBLIC PERCEPTIONS

Vaccine safety assessment is multifaceted, involving review of manufacturing processes, animal studies, clinical trials, and postlicensure surveillance.[3] In phase 1 studies, involving approximately 20 to 80 participants, participants are closely monitored for adverse events (e.g., fever, injection site reactions). In the case of live viral and bacterial vaccines, investigators look for evidence of shedding and level of attenuation, among other issues. Vaccine-elicited immune responses are also assessed.

Phase 2 trials are often randomized, well-controlled studies that enroll up to several hundred individuals to assess immunogenicity and any other markers of potential efficacy. These trials offer opportunities to compare rates of common local and systemic reactions in groups assigned to different dose levels or a placebo. Serious concerns can sometimes be documented in these early studies. For example, early in its development in the 1960s, an inactivated respiratory syncytial virus vaccine was shown to enhance illness severity.[4]

Phase 3 trials, which provide the primary evidence of efficacy and extend the safety database, generally involve randomization of participants to the new vaccine or a placebo or alternative vaccine in a blinded fashion. Randomization and blinding are necessary to control bias in regard to participant entry, treatment assignment, management, and

follow-up so that safety as well as efficacy outcomes are readily interpretable. Simplified study designs, with active monitoring of common events only in a relatively small patient subset, have been used when efficacy assessment has required very large studies.[5,6] The largest vaccine trial ever performed—perhaps the largest trial of any medical intervention ever performed, and almost surely the simplest—was the mid-1950s Francis et al. field trial of the Salk inactivated polio vaccine.[7] This trial involved randomization of more than 400 000 children; cases of paralytic polio were identified from reports sent to local health departments rather than through individualized follow-ups.[8]

There are no published guidelines regarding the size of vaccine trials. In the case of new combination vaccines consisting of licensed components previously subject to safety evaluation, the FDA generally advises sponsors that they need a minimum of about 5000 participants (vaccinees in addition to control participants) to allow the necessary safety and immunogenicity evaluation in randomized controlled trials. When a new single vaccine or vaccine component is being studied, the number of participants may be much larger, depending on the particular endpoint chosen and the incidence of the disease to be prevented. In some cases, a larger study may be required because of a need for information about specific safety parameters. Often, manufacturers are asked for commitments to accumulate additional postlicensure safety data, but such studies are seldom randomized and may be uncontrolled. Occasionally, postlicensure registries are established for specific concerns such as vaccination during pregnancy, wherein the objective is to identify any possible adverse effect on the fetus.

Although vaccines in current use have almost eliminated mortality and morbidity resulting from polio, measles, diphtheria, tetanus, and multiple other diseases in the United States, continuing and ongoing postlicensure monitoring of very rare serious events will always be needed. The "Cutter incident," in which cases of paralytic polio resulted from administration of polio vaccine that had not been successfully inactivated, demonstrated early on the critical role of ongoing safety surveillance for new vaccines.[9] Other early safety concerns discovered via postlicensure surveillance included paralytic poliomyelitis associated with oral polio vaccine,[10] complications of smallpox vaccination,[11] and Guillain-Barré syndrome associated with influenza vaccine.[12]

The National Childhood Vaccine Injury Act of 1986 requires health professionals and vaccine manufacturers to report specific adverse events that occur after the administration of routinely recommended vaccines to the US Department of Health and Human Services (42 USC §300aa-25). In 1990, as a result of this legislation, the CDC and FDA jointly initiated the Vaccine Adverse Event Reporting System (VAERS) to provide a single system as a replacement for the separate monitoring programs operated by the 2 agencies.[13]

The goals of safety surveillance, in addition to early detection of potential rare but serious vaccine reactions, include ensuring public confidence in the safety, value, and importance of immunization. Multiple approaches are needed to achieve surveillance goals. Passive surveillance systems such as VAERS have important strengths but also major limitations.[14–16] The strengths of VAERS include its national scope (any member of the public may report adverse events to VAERS), potentially timely reporting, ability to identify extremely rare events, and ability to detect individual vaccine lots with unusual reporting patterns.

VAERS has many serious limitations, however, including underreporting, inadequate or inaccurate data provided by reporters, lack of the denominator data needed to estimate rates, and lack of known background rates with which to compare VAERS reporting rates (number of reports received divided by number of doses distributed). It is also difficult to interpret reported adverse events in the presence of multiple simultaneous vaccinations. Thus, VAERS data can only rarely support causality assessments; they are most useful for identifying signals that may trigger more formal and rigorous investigation.

Another important and more recently instituted tool for postlicensure vaccine safety surveillance is the Vaccine Safety Datalink (VSD).[17] The VSD, a consortium of health maintenance organizations (HMOs), was established in 1991 to provide information on important medical outcomes with linkages to vaccination records in a defined population. In addition, key information such as demographic and health-related characteristics and birth and death certificate data are available. Further information can be derived from medical chart reviews, facilitating the conduct of epidemiological investigations such as case–control studies.

However, the VSD also has weaknesses, including its limited sample size, limited population diversity, and very small (and surely unrepresentative) unvaccinated population. VSD investigations have followed up on VAERS signals and on specific public concerns. Many issues, such as the possible relation of autism to receipt of thimerosal-containing vaccines, diabetes mellitus following *Haemophilus influenzae* type b (Hib) conjugate or hepatitis B vaccination, and chronic arthropathy among women receiving rubella vaccine, have been studied through the VSD; no evidence of increased risk for these conditions has been found in association with vaccination.[18–20] The VSD has been expanded, and the CDC is working with the Institute of Medicine to set research priorities.

Additional approaches to improving vaccine safety surveillance could be considered, but pragmatic issues must be taken into account. Such considerations might include availability of exposure and outcome data, resources for data management, infrastructure to support data retrieval and reporting, and extent of impact on the daily activities of health care providers. Potential vehicles for capturing data include other HMO research networks and the American Academy of Pediatrics Practice Research Office Settings (PROS).[21] PROS involves many of the attributes just noted, but vaccine usage is not tracked in any standardized way. Creation of a new data infrastructure would require substantial resources, and one would have to consider whether participating providers would accept the additional reporting requirements; whether timely data could be made available through this system and, critically, whether it could be implemented without generating an overwhelming number of false signals; and whether another vaccine safety system represents the best use of the resources that would be required.

Parental concerns about vaccine safety play an increasingly important role in public health activities and policy setting.[22] Issues affecting risk perception include ability to control exposure, whether effects are immediate or delayed, reversibility of effects, level of trust in responsible institutions, and media attention. The primary sources of public information on vaccine safety are physicians (especially pediatricians), parenting books and magazines, the Internet, and friends in health care–related fields. A recent poll conducted by several professional and public health organizations indicated a "simmering" worry about vaccine safety.[23]

SESSION 2: RATIONALE FOR LARGER PRELICENSURE CLINICAL TRIALS

Vaccines are different from most other pharmaceuticals in ways that influence safety considerations. They are administered to millions of healthy people every year, and, unlike any therapeutic product, many vaccines are mandated for entry into schools, preschool programs, and day care programs by most states, as well as during military service. Vaccines are known to cause a small number of extremely rare but serious adverse outcomes, such as vaccine-associated paralytic polio and anaphylactic shock, but are suspected by some to be responsible for a much larger number and wide variety of serious and chronic health problems. It is often difficult to address such concerns effectively because, as discussed earlier, evaluating vaccine safety in the postlicensure setting is very complicated. With vaccines, as with other widely used products, some serious medical events will occur coincidentally after administration. Often it is impossible at that time to ascertain the likelihood of any causal connection with the vaccine; because such events occur in the absence as well as the presence of vaccines, causal effect is rarely demonstrable.

Controlled trials conducted prior to licensure are designed to establish efficacy and assess relatively common adverse events, and they are usually far too small to detect rare, serious outcomes that could still affect large numbers of children each year. A trial of 5000 participants, equally divided between new vaccine and control groups, would have good power to detect a doubling of an adverse event that might occur in 1% of the population but would have virtually no power to detect a doubling of an event occurring in only 0.1% of the population. Such an adverse event would, however, affect as many as 4000 children a year in the United States alone.[24] Also, the inclusion criteria in these trials are often narrow, excluding children with chronic or acute illness; broadening these criteria would allow some data to be obtained, in a controlled setting incorporating active monitoring for safety, with children who might be at increased risk of adverse events and ultimately would be part of the vaccine's target population.

The concept of routinely enrolling tens of thousands of individuals in controlled trials of new vaccines for the purpose of obtaining better prelicensure safety information, even when such large trials are not necessary for determination of efficacy, is controversial. Major concerns include the expense of such trials (possibly creating additional disincentives for vaccine development), the potential delayed availability of important new vaccines, and the appearance of more spurious associations with adverse outcomes that would raise unwarranted safety concerns. Furthermore, workshop participants recognized that larger trials would not detect all problems; increases in adverse events occurring at background rates of 0.01% or less would not likely be detected in trials of feasible size, and the problem of identifying latent effects that might not be identified for months or years after vaccination would remain.

These concerns, while valid, provide challenges but perhaps not insurmountable obstacles to the conduct of large controlled trials of some new vaccines. Spurious associations with adverse outcomes must be addressed in trials of any size; although larger trials will allow observation of less common events whose association with vaccination may be spurious, the ability to exclude the possibility of major risks of serious outcomes would be greatly enhanced by larger numbers. Increased cost is certainly an issue, but there are many ways to increase trial efficiency that could be implemented and that have, in fact, been implemented in trials in a variety of medical settings, including vaccines.

Simplified procedures that facilitate inclusion of a large number of sites might permit larger trials to be conducted with little or no additional time requirements. The issue of late-occurring events could be addressed by extending follow-up of trial participants to a year (or more, if needed) after their most recent vaccination, possibly as a postmarketing commitment to avoid delaying vaccine availability. Finally, the argument that it is not feasible to perform trials large enough to detect or rule out the rarest adverse events does not seem to justify avoiding the question of where the line should be drawn: Is it sufficient to detect substantial increases only of adverse events occurring at rates of 1% or more, or should the safety standard for new vaccines be more stringent?

Large trials have been carried out in other areas such as cardiovascular disease and cancer. Researchers working in these areas have been motivated by the recognition that many treatments in these settings became widely used without being reliably assessed and that effects of treatments are likely to be of modest size and not detectable in small trials.[25] Reliable results depend on minimizing both biases (accomplished by randomization and blinding) and random errors (accomplished by large sample sizes). It is neither possible nor necessary to completely eliminate biases or random error. It is critical, however, to ensure that residual biases and random errors are small relative to the size of the effect to be measured. Even incomplete information can contribute to valid conclusions as long as it is unbiased. As an example, a study conducted in China in which approximately 70 000 children were randomized (in a cluster randomization design) to either a group receiving hepatitis B vaccine or a control group relied on a cancer registry for long-term follow-up information on liver cancer rather than decades-long individual follow-up of vaccinees.[26]

Staggered introductions of new vaccines may also provide useful safety information when, for administrative reasons, vaccines cannot be introduced universally and simultaneously. Rather than implementing vaccination programs haphazardly in different places, one could consider introducing the new vaccine in a random way, facilitating

objective and unbiased assessments of effects.

Experience has shown that studies are easiest to conduct when they involve very nonrestrictive eligibility criteria, consistent with the expected target population, so that little effort is needed to assess eligibility. Streamlined trial entry procedures and minimal data collection facilitate conducting trials in busy clinics.

There are many ways to simplify trial conduct. Simplified procedures for case ascertainment, as in the hepatitis B vaccine trial just discussed, may be acceptable, even if some cases are missed, as long as there is no bias. Placebo-controlled trials, which require twice as many injections as trials involving controls who are simply unvaccinated (assuming a 1:1 randomization schedule), may not be necessary if the ascertainment of efficacy and safety outcomes is unbiased. Data on common adverse events and immune responses could be collected on a small subset of participants, with minimal data collection on the majority of the study population. Cluster randomized trials, in which clinics rather than individuals are randomized,[27] are especially simple to conduct because of the reduced complexity associated with clinic operations.

Simplified randomized trials will produce results far more reliable than those derived from observational postmarketing surveillance efforts because random assignment minimizes concerns about confounding and other problems associated with observational data.[28] In the case of many of the rare events of interest, only large-scale, well-designed randomized trials can provide the security of knowing that one can rule out a high risk of important adverse effects.

SESSIONS 1 AND 2: DISCUSSION

The concept of large, simple trials to evaluate new vaccines raised many questions among the workshop attendees. The first was whether trials of 50 000 individuals would actually be adequate to detect the sorts of problems that have been at the forefront of public concerns regarding vaccines—autism, for example. Other concerns were related to availability of sufficient trial sites, potential effects of simplified trial procedures on quality control, extent to which the seriousness of the disease to be prevented should influence the required size of a safety database, and the possibility that demands for larger trials could create a major disincentive for new vaccine development.

Some of the workshop participants expressed concern that results from trials involving heterogeneous populations would not be readily interpretable; others noted the advantages of studying a population that was more representative of the ultimate target population. The inferential problem of considering multiple types of adverse events that could not always be prespecified was raised and acknowledged; as noted earlier, this problem exists for small trials as well as large trials. The importance of randomized comparisons in assessing relative rates of events, and recognizing elevated rates, was emphasized. There was limited discussion of incremental (rather than immediate universal) introduction of new vaccines, which would provide some opportunity to collect comparative safety data, albeit not randomized, as an alternative to large randomized trials. There was general agreement that a robust postlicensure surveillance system would always be needed, regardless of the size of prelicensure trials.

SESSION 3: EXPERIENCE WITH LARGE TRIALS

The Northern California Kaiser Permanente efficacy trial of the heptavalent pneumococcal conjugate vaccine included approximately 38 000 infants and was the primary basis for licensure of this vaccine in 2000.[5] The study was large because the efficacy measure—invasive pneumococcal disease—was relatively rare, but the large numbers offered the opportunity to evaluate a considerable amount of safety data. Some aspects of study conduct were streamlined in this large trial, but quality control and adherence to study procedures were well maintained. An automated tracking system captured diagnoses, procedures, and laboratory test results and was used to identify uncommon adverse events. In addition, telephone interviews were conducted with some of the parents to document local and systemic reactions. The infrastructure provided by the Kaiser system, with automated availability of clinical data, helped to ensure consistency and quality control.

Large trials of acellular pertussis (aP) vaccines were conducted in Sweden during the 1980s and 1990s.[29,30] Because of concerns about the safety of whole cell pertussis vaccines, an important goal of these trials was to accumulate substantial safety data before introduction of an aP vaccine (in combination with diphtheria and tetanus toxoid vaccines [i.e., DTaP]) to the entire population. One trial of approximately 10 000 infants (trial 1) compared DTaP vaccines containing 2 and 5 aP components, respectively, with a US-licensed whole cell DTP vaccine as well as a DT only (no pertussis component) vaccine.[29] Trial 2 was much larger (involving 82 000 infants) and compared DTaP vaccines containing 2, 3, and 5 aP components with a British whole cell DTP vaccine.[30] Because trial 1 enrollment was completed before trial 2 was initiated, and an interim analysis had shown acceptable safety, trial 2 relied on passive surveillance for adverse events, a substantial simplification. More than 80% of the eligible cohort participated in trial 2, owing in part to a resurgence of pertussis in Sweden during the trial.

Much of the critical effort in trials such as these occurs before the trial, in the development of the protocol, parental consent information, concise and targeted data collection forms, the analysis plan, and training of participating investigators and staff. In well-designed, randomized trials, researchers can examine long-term consequences of vaccination using national health registries (where they exist) to learn about hospitalizations and using other resources (such as the Swedish child diabetes registry[31]) to investigate chronic health problems. Such trials can also provide opportunities to reliably address questions that arise later.

A controlled trial of a Hib conjugate vaccine conducted in Finland in the mid-1980s included more than 114 000 children enrolled at about 1000 different sites.[6] Children were assigned to receive Hib conjugate vaccine either in infancy (3, 4, 6, or 14–18 months) or at the age of 24 months, according to birth date. Active safety surveillance for common adverse events was performed among a small subset of 99 children; investi-

gators relied on passive surveillance for identifying potential serious adverse events. Investigators were encouraged to report serious adverse events both to the trial group and to the National Board of Health. This large trial showed that the vaccine was effective and led to its universal use (and ultimately to the virtual elimination of invasive Hib disease) in Finland. The database has continued to be a valuable resource for further studies of possible rare vaccine risks such as type 1 diabetes[32] and childhood leukemia.[33]

A trial comparing ibuprofen with acetaminophen for treatment of fever in children is a particularly interesting example of how a large trial might be done simply. The trial was designed to assess the relative safety of ibuprofen, a nonsteroidal anti-inflammatory drug (NSAID), when it was proposed for consideration as an "over-the-counter" drug for children.[34] At the time, there was relatively little pediatric experience with ibuprofen, and concerns about the safety of antipyretics had been heightened by the association of aspirin with Reye's syndrome. An observational study, however, was thought to be insufficiently reliable for assessment of possible excess risks of ibuprofen; thus, despite the additional complexity, a randomized controlled trial was initiated in primary care practice settings. The study design had to be uncomplicated so that it did not disrupt office routines, yet sufficiently rigorous to provide credible results. Children were randomly assigned to 1 of 2 different doses of ibuprofen or to acetaminophen. Primary care physicians were enthusiastic about the trial because the question addressed was clinically relevant to them and because all children received an antipyretic (i.e., there was no placebo arm).

Primary study outcomes were hospitalizations for events known to be associated with NSAID use in adults: acute gastrointestinal bleeding, renal failure, and anaphylaxis. Reye's syndrome was also included as a primary outcome because of its known association with aspirin. Outcomes were assessed via questionnaires sent to parents and telephone interviews of parents who did not return questionnaires. Medical records were obtained for reported hospitalizations. The final study report was based on 83 915 children enrolled by 1735 physicians over 29 months. No differences in the outcomes of interest were observed. No cases of renal failure, anaphylaxis, or Reye's syndrome were reported; the large sample size allowed the conclusion that such events would occur at a rate of no greater than 5 in 100 000 children treated with ibuprofen.

Large trials such as this may be "simple" in some sense, but they are difficult to organize. However, they can be feasible if the hypothesis is clinically relevant to practicing physicians and if the study procedures are simplified so that intrusions on the conduct of medicine in outpatient practices are minimized.

As the final report for the ibuprofen study was being prepared, several pediatric cases of necrotizing fasciitis (external to this trial) were reported, with a high rate of NSAID exposure prior to hospitalization. The database from this large controlled trial offered the opportunity to assess any possible associations of either treatment with this disease; no such associations were identified.

A final example focused on design considerations for a study (now completed) of a human–bovine reassortant rotavirus vaccine. A key feature of the study was its attention to any possible association with intussusception, an uncommon adverse event associated with the rhesus rotavirus vaccine licensed earlier but then withdrawn. Intussusception is extremely rare among infants younger than 2 months; the rate increases somewhat in older infants. Rotavirus vaccine was administered at 2, 4, and 6 months of age; thus, during this interval, the expected background rate of the adverse event is increasing, but the event remains rare.

Designing and implementing a trial to evaluate efficacy as well as occurrence of a relatively rare adverse event is challenging. Clearly, the trial must provide a high level of confidence that the vaccine is acceptably safe. Because intussusception has potential complications and sequelae that can be serious or even fatal, the trial must minimize the risk to participants. From the sponsor's perspective, the potential for a result that will support acceptance on the part of both regulatory agencies and the medical community must be sufficient to justify the expense and effort of the trial. Because ongoing safety monitoring is imperative, the trial design must allow for early termination if interim evidence shows that the vaccine does in fact increase the rate of intussusception (i.e., the observed data cross a predefined safety boundary). The randomized, double-blind, placebo-controlled trial (which enrolled more than 70 000 participants) was designed to ensure a sound statistical basis for the intussusception safety evaluation, with intensive interim review by an independent committee and an endpoint adjudication committee to assess all reports of intussusception.

SESSION 3: DISCUSSION

Four scenarios in new vaccine development were discussed with regard to whether they might warrant large, prelicensure trials with a focus on safety issues: (1) the first vaccine to address a particular disease, (2) any new pediatric vaccine, (3) a vaccine likely to be recommended for universal use, and (4) a vaccine (or class of vaccines) for which a prior safety issue has already been identified. No consensus was reached on any of the first 3 scenarios; there was general agreement that large trials focusing on safety issues might be warranted in the fourth scenario, depending on the severity and frequency of the adverse event in question.

Some participants noted that large trials may be less important for a vaccine that is substantially similar to existing vaccines, although criteria for establishing "similarity" may not always be clear. There was general agreement that a new combination vaccine for which there has been substantial experience with the vaccine components would raise fewer safety concerns than an entirely new vaccine. The severity of the target disease and the size of the target population are relevant to the need for a large randomized safety database. Many workshop participants believed that the success of the current system in rapidly detecting the association of intussusception with rotavirus vaccine showed that this system is working well. Concerns were expressed about interpreting results of large safety trials without specific a priori safety hypotheses, although the need to conduct exploratory safety analyses in any study, whatever its size, was noted.

The "natural experiment" that occurs immediately after the introduction of a new vaccine, when use of the vaccine is gradually increasing, was discussed as a potential source of data on vaccine safety. The concept of a deliberate staged introduction of new vaccines, however, generated the concern that random provision of an approved vaccine to some population segments before others might require informed consent from the entire population. The possibility of establishing a control group of children whose parents would prefer to wait for more data before having their children vaccinated was raised. This approach would take advantage of differing perspectives on potential benefits and risks but could be subject to severe selection bias. When there is no clear a priori hypothesis, safety endpoints for large trials would be those serious events that could be most reliably captured (e.g., those requiring hospital visits or visits to emergency rooms or physicians' offices in settings in which data from such visits are available). Information obtained from parents via mail or telephone could also be assessed but could not be documented as well. The ability to assess a potentially causal association of a vaccine with a subsequent adverse event is clearly greatest in the context of a randomized trial.

Another possible approach to generating large, well-controlled safety databases would be to initiate controlled safety studies immediately following completion of the efficacy trial (or trials) but before regulatory approval or recommendations for universal use, or to simply continue to randomize participants during the regulatory evaluation process. A problem with conducting controlled trials after licensure but before recommendations for use relates to diversity in access to health care; if the vaccine could not be recommended for universal use until the large safety study was completed, there would be a period during which people who could afford the vaccine could have their children protected, while those who relied on insurance coverage or public clinics would not have access to the vaccine. Another problem would be potential resistance to performing controlled studies of an already-licensed vaccine, even one not yet recommended for universal use. Conducting all randomized controlled studies before licensure, including, possibly, a large "postefficacy demonstration" safety trial, might avoid these problems.

The potential of epidemiological studies to identify important safety issues cannot be discounted. Improved methods for studying risks in observational studies, such as considering events occurring within and outside predefined "risk windows" following vaccination, have enhanced the likelihood that observational studies will provide useful information. The reliability of such methods is questionable, however, when the relative risks are modest or moderate (i.e., no more than twofold or threefold). Workshop participants noted that concerns about the safety of exposing large numbers of children to an investigational vaccine must be considered in light of the fact that, after licensure, millions of children may be exposed to the vaccine with no oversight other than that afforded by passive surveillance systems. In a large trial, with careful monitoring by the sponsor, investigators, and usually an independent data monitoring committee, an unexpected harmful effect is more likely to be identified, diagnosed, and considered as a possible vaccine reaction, thereby protecting future vaccine recipients and contributing information about the safety of the vaccine.

The use of large trial databases to explore safety questions emerging later (as in the case of some of the trials described previously) is appealing, although confidentiality concerns could raise obstacles. Also, such analyses would be partially confounded if (as often occurs) individuals who received placebos in these trials were provided the vaccine at trial termination and if age at exposure were not correlated with the outcome of interest.

SESSION 4: IMPLICATIONS OF LARGE TRIALS

From the industry's perspective, it is good business to have a safe vaccine; even the perception of a safety problem can have serious negative consequences. Large trials may provide useful information but can be very costly to conduct and could delay time to market. It is not clear how much society is willing to pay for the extent of risk reduction that might be afforded by larger trials. There are many other ways to potentially improve the safety of vaccines; for example, improved purification in the manufacturing process could eliminate extraneous materials that might cause adverse events. Unfortunately, vaccines are undervalued by society at large, by physicians, by health planners, and by policymakers both in the United States and elsewhere, even though vaccination is one of the relatively few modalities providing a clear cost savings to society in terms of reduced medical expenditures alone (i.e., not including time lost from work and other relevant considerations).

Clearly, not all types of problems will be detected with a large trial. For example, despite the inclusion of more than 1 million children in the Salk inactivated polio vaccine trial, postlicensure manufacturing and quality control problems led to the "Cutter incident."[9] In addition, contamination of early vaccine products with a macaque polyomavirus (SV-40) that was present in macaque kidney tissue used in the manufacture of inactivated polio vaccine was not recognized until years after the vaccines were put into use.[35,36]

Although larger prelicensure trials would improve the likelihood of detecting less common adverse events (and ruling out vaccine associations with such events at a more precise level), assessment of extremely rare events, late occurring adverse events, and adverse events occurring in selected small subgroups (e.g., premature infants) would still be problematic. Public confidence might be enhanced by the perception that adverse reactions would more likely be detected, but it could also be damaged by unwarranted concerns about the spurious associations that inevitably arise or even about low-incidence, actual associations that are observed but are of questionable public health significance.

The possible consequences of larger prelicensure trials that have been discussed—delay of vaccine licensure and widespread use and increased costs of bringing a product to licensure—could have an adverse impact on public health programs. Alternative approaches to CDC's recommendation process could be used to collect additional safety data without unduly delaying the availability of important new vaccines; for example, initial

recommendations for vaccine use could be limited and extensive review of subsequent postlicensure data required prior to expanding recommendations.

GENERAL DISCUSSION

Some meeting participants were concerned that "raising the bar too high" might adversely affect future vaccine development. Further exploration of the use of HMO settings for vaccine trials was advocated, given that information on primary outcomes and serious adverse events is already collected in an easily accessible electronic database. This approach would provide broader experience as a basis for use recommendations and might enhance public acceptance of a vaccine. It was noted that more large trials of new drugs have been implemented in recent years and that manufacturers have learned how to perform them more efficiently. Many participants expressed the concern that requiring larger trials would necessarily delay licensure of effective vaccines, thus raising costs and delaying availability of the benefits such vaccines would offer. There was some skepticism about the value of the gain in safety information that larger trials would provide, relative to the additional costs and delays, but the potential for long-term gains in acceptance of vaccines if risks are assessed more carefully was also noted.

Most participants realized the desirability of a less abrupt transition between the thousands of people vaccinated in prelicensure trials to the millions vaccinated postlicensure. In this regard, participants proposed delaying vaccine use recommendations until more safety data are available in the immediate postlicensure period. There was little disagreement with the principle that different approaches would probably be needed for different vaccines.

It was further proposed that a standard be set for minimum sample sizes in prelicensure randomized, controlled trials of any new vaccine (with new antigens) for which universal use is likely to be recommended and that 10 000 participants (vaccinees in combination with controls) might be the appropriate standard; sample sizes in studies focusing on vaccines used in the prevention of milder diseases might be higher. Many other participants agreed with the concept of a "floor" value for vaccine trial sizes, but consensus regarding the basis for selecting that value did not emerge.

The notion of "phasing in" a new vaccine, with initial administration to populations at greatest risk from the targeted disease, was again advocated, especially with provisions for active data collection on safety outcomes during the initial period of use. Large, controlled safety trials would be implemented after the initial, more intensive efficacy trial; these trials could be substantially simplified, relying on postcard reports or even Internet reporting to collect data only on major events. Vaccine developers would need to implement such safety trials shortly after the completion of efficacy trials. Participants noted that often there is no easy way to identify subpopulations at different levels of risk; they further noted that the logistical difficulties of enrolling very large numbers of healthy infants should not be underestimated. There was concern that although active postlicensure data collection is feasible, postlicensure conduct of randomized controlled trials (especially placebo-controlled trials) might raise ethical concerns. The possibility was raised of having an independent oversight group, along the lines of a data monitoring committee, evaluate accumulating safety data in the initial postlicensure period.

Another of the meeting participants, a parent of a child who had contracted polio from the oral polio vaccine, indicated his belief that if the public were more aware of the extent of efforts to ensure vaccine safety, even at the current levels, the comfort level with vaccination would increase. He also expressed his belief that, despite the importance of preventing serious diseases such as polio, more attention should be focused on vaccine safety as well as efficacy.

Participants noted that more direct approaches to improving vaccine safety might be possible. There may be ways to develop less reactogenic vaccines based on the known pathogenesis of adverse reactions. Increasing the purity of vaccines or new subunit vaccines might reduce local reactions. Some suggested that recent technology making it possible to fill vaccines and maintain sterility without using bacteriological agents such as antibiotics could help avert possible rare reactions to an antibiotic. Such improvements, if recognized by the public, might increase public confidence. Another approach, developing the ability to identify individuals with possible genetic predispositions for adverse reactions to a vaccine so that they could avoid vaccination, has been advocated; however, few such genetic susceptibilities are known. Furthermore, identifying and validating these susceptibilities and then implementing widespread screening programs would be extremely difficult. In addition, such approaches would probably not be sufficient to address most potential safety issues.

CONCLUSIONS

Overall, there appeared to be general agreement among the workshop participants that different vaccines would present different safety considerations. In some situations, the current approach of moving from relatively small data sets in prelicensure studies to immediate universal vaccination should probably be reconsidered. Alternatives discussed for such situations included (1) expanding prelicensure trials; (2) initiating, after completion of traditionally sized efficacy trials, large controlled safety trials that would be conducted during the time the primary data were under regulatory review and whose data would be available within the early postmarketing period; and (3) phasing in vaccine use, with perhaps more active surveillance of safety data, for a certain period of time after licensure but prior to any recommendation for universal vaccination.

There was also substantial support for setting a minimum sample size for prelicensure trials of vaccines likely to be recommended for universal use. The current system produces vaccines that are highly effective and safe, but improvements may be necessary to counter increasing public concerns. Consideration of requirements for safety evaluation will need to be made case by case, according to the size and characteristics of the target population, the severity and overall burden of the targeted disease, and other factors. The more the public recognizes the efforts of vaccine manufacturers and public health agen-

cies to minimize vaccine risks, the more confidence the public will have in the safety of vaccines. ■

About the Authors

At the time this article was written, Susan S. Ellenberg was with the Center for Biologics Evaluation and Research, US Food and Drug Administration, Rockville, Md. Mary A. Foulkes, Karen Midthun, and Karen L. Goldenthal are with the Center for Biologics Evaluation and Research, US Food and Drug Administration.

Requests for reprints should be sent to Susan S. Ellenberg, PhD, University of Pennsylvania School of Medicine, Center for Clinical Epidemiology and Biostatistics, Division of Biostatistics, Blockley Hall, Room 611, 423 Guardian Dr, Philadelphia, PA 19104-6021 (email: sellenbe@cceb.upenn.edu).

This article was accepted August 8, 2004.

Contributors

S.S. Ellenberg and M.A. Foulkes prepared the initial drafts of the article. K. Midthun and K.L. Goldenthal participated in the extensive editing process and rewrote major sections of the article.

Acknowledgments

The authors thank Dr Carole Heilman and the late Dr John LaMontagne for contributing to the organization of the workshop, as well as all of the workshop participants for providing the important perspectives and ideas summarized in this article.

Workshop speakers and panelists (in order of presentation) were as follows: Kathryn Zoon, PhD, Karen L. Goldenthal, MD, Walter Orenstein, MD, MPH, Bruce Gellin, MD, Robert Davis, MD, Susan S. Ellenberg, PhD, Rory Collins, MD, Steven Black, MD, Patrick Olin, MD, Juhani Eskola, MD, Samuel Lesko, MD, Jerry Sadoff, MD, Thomas Fleming, PhD, Ted Eickhoff, MD, William Blackwelder, PhD, Gordon Douglas, MD, Peter Patriarca, MD, MPH, George Siber, MD, Robert Temple, MD, Karen Midthun, MD, Benjamin Schwartz, MD, MPH, Norman Fost, MD, Jesse Goodman, MD, MPH, Neal Halsey, MD, Samuel Katz, MD, Randy Pearson, and John LaMontagne, PhD.

References

1. Centers for Disease Control and Prevention. Withdrawal of rotavirus vaccine recommendation. *MMWR Morb Mortal Wkly Rep.* 1999;48:1007.

2. Murphy TV, Gargiullo PM, Massoudi MS, et al. Intussusception among infants given an oral rotavirus vaccine. *N Engl J Med.* 2001;344:564–572.

3. Goldenthal KL, McVittie L. The clinical testing of preventive vaccines for infectious disease indications. In: Mathieu M, ed. *Biologics Development: A Regulatory Overview.* 2nd ed. Waltham, Mass: Parexel; 1997:123–139.

4. Kim HW, Canchola JG, Brandt CD, et al. Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. *Am J Epidemiol.* 1969;89:422–434.

5. Black S, Shinefield H, Fireman B, et al. Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. *Pediatr Infect Dis J.* 2000;19:187–195.

6. Eskola J, Kayhty H, Takala AK, et al. A randomized, prospective field trial of a conjugate vaccine in the protection of infants and young children against invasive *Haemophilus influenzae* type b disease. *N Engl J Med.* 1990;323:1381–1387.

7. Francis TM Jr, Korns RF, Voight RB, et al. An evaluation of the 1954 poliomyelitis vaccine trials (summary report). *Am J Public Health.* 1955;45:1–51.

8. Meier P. Polio trial: an early efficient clinical trial. *Stat Med.* 1990;9:13–16.

9. Nathanson N, Langmuir AD. The Cutter incident: poliomyelitis following formaldehyde-inactivated poliovirus vaccination in the United States during the spring of 1955. II. Relationship of poliomyelitis to Cutter vaccine. *Am J Epidemiol.* 1995;142:109–140.

10. Strebel PM, Sutter RW, Cochi SL, et al. Epidemiology of poliomyelitis in the United States one decade after the last reported case of indigenous wild virus-associated disease. *Clin Infect Dis.* 1992;14:568–579.

11. Goldstein JA, Neff JM, Lane JM, Koplan JP. Smallpox vaccination reactions, prophylaxis, and therapy of complications. *Pediatrics.* 1975;55:342–347.

12. Lasky T, Terracciano GJ, Magder L, et al. The Guillain-Barré syndrome and the 1992–1993 and 1993–1994 influenza vaccines. *N Engl J Med.* 1998; 339:1797–1802.

13. Chen R, Rastogi SC, Mullen JR, et al. The Vaccine Adverse Event Reporting System (VAERS). *Vaccine.* 1994;12:542–550.

14. Ellenberg SS, Chen RT. The complicated task of monitoring vaccine safety. *Public Health Rep.* 1997;112: 10–20.

15. Chen RT. Safety of vaccines. In: Plotkin S, Orenstein W, eds. *Vaccines.* 3rd ed. Philadelphia, Pa: WB Saunders Co; 1999:1144–1163.

16. Ellenberg SS, Braun MM. Monitoring the safety of vaccines: assessing the risks. *Drug Saf.* 2002;25: 145–152.

17. Chen RT, Glasser JW, Rhodes PH, et al. Vaccine Safety Datalink project: a new tool for improving vaccine safety monitoring in the United States. *Pediatrics.* 1997;99:765–773.

18. Verstraeten T, Davis RL, DeStefano F, et al. Safety of thimerosal-containing vaccines: a two-phased study of computerized health maintenance organization databases. *Pediatrics.* 2003;112:1039–1048.

19. DeStefano F, Mullooly JP, Okoro CA, et al. Childhood vaccinations, vaccination timing, and risk of type 1 diabetes mellitus. *Pediatrics.* 2001;108:E112.

20. Ray P, Black S, Shinefield H, et al. Risk of chronic arthropathy among women after rubella vaccination. *JAMA.* 1997;278:551–556.

21. Wasserman RC, Slora EJ, Bocian AB, et al. Pediatric research in office settings (PROS): A national practice-based research network to improve children's health care. *Pediatrics.* 1998;102:1350–1357.

22. Wilson CB, Marcuse EK. Vaccine safety–vaccine benefits: science and the public's perception. *Natl Rev Immunol.* 2001;1:160–165.

23. Gellin BG, Maibach EW, Marcuse EK. Do parents understand immunizations? A national telephone survey. *Pediatrics.* 2000;106:1097–1102.

24. Ellenberg SS. Safety considerations for new vaccine development. *Pharmacoepidemiol Drug Saf.* 2001; 10:411–415.

25. Yusuf S, Collins R, Peto R. Why do we need some large, simple randomized trials? *Stat Med.* 1984;3: 409–422.

26. Sun Z, Zhu Y, Stjernsward J, et al. Design and compliance of HBV vaccination trial on newborns to prevent hepatocellular carcinoma and 5-year results of its pilot study. *Cancer Detect Prev.* 1991;15:313–318.

27. Donner A, Klar NS. *Design and Analysis of Cluster Randomization Trials.* London, England: Hodder Arnold; 2000.

28. Hill AB. *Principles of Medical Statistics.* 9th ed. New York, NY: Oxford University Press Inc; 1971.

29. Gustafsson L, Hallander HO, Olin P, et al. A controlled trial of a two-component acellular, a five-component acellular, and a whole-cell pertussis vaccine. *N Engl J Med.* 1996;334:349–355.

30. Olin P, Rasmussen F, Gustafsson L, et al. Randomised controlled trial of two-component, three-component, and five-component acellular pertussis vaccines compared with whole-cell pertussis vaccine. *Lancet.* 1997; 350:1569–1577.

31. Nystrom L, Dahlquist G, Rewers M, Wall S. The Swedish Childhood Diabetes Study: an analysis of the temporal variation in diabetes incidence 1978–1987. *Int J Epidemiol.* 1990;19:141–146.

32. Karvonen M, Cepaitis Z, Tuomilehto J. Association between type 1 diabetes and *Haemophilus influenzae* type b vaccination: birth cohort study. *BMJ.* 1999;318: 1169–1172.

33. Auvinen A, Hakulinen T, Groves F. *Haemophilus influenzae* type b vaccination and risk of childhood leukaemia in a vaccine trial in Finland. *Br J Cancer.* 2000;83:956–958.

34. Lesko SM, Mitchell AA. An assessment of the safety of pediatric ibuprofen: a practitioner-based randomized clinical trial. *JAMA.* 1995;273:929–933.

35. Lewis AM, Egan W. Workshop on simian virus 40 (SV40): a possible human polyomavirus. *Biologicals.* 1997;25:355–358.

36. Butel JS. Simian virus 40. In: Webster RG, Granoff A, eds. *Encyclopedia of Virology.* Vol. 3. San Diego, Calif: Academic Press; 1994:1322–1329.