Invited Commentary: Use of Selective Viral Cultures to Adjust Nonvirologic Endpoints in Studies of Influenza Vaccine Efficacy

Arnold S. Monto

From the Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI.

Received for publication April 17, 2003; accepted for publication May 22, 2003.

Inactivated vaccine: endpoints used to determine efficacies

Influenza vaccine was first developed during the Second World War to protect the US military and to maintain readiness of combat troops (1). The virus was grown in fertile eggs, and the antigen was concentrated and inactivated before administration. Vaccines currently licensed for use in the United States are produced by using similar techniques. They are now more highly standardized, and the virus has been disrupted so that viral components considered unnecessary for protection can be removed; in general, however, they are little changed from those developed approximately 60 years ago (2). Most estimates of protective efficacy are also old. Nearly every year through the 1960s, randomized controlled trials comparing vaccine with placebo were conducted in the US military. The endpoint for comparison was infection as determined not by virus isolation but by rise in antibody titer, based on the hemagglutination-inhibition test, between the blood specimens collected before and after the influenza season. Protective efficacy, so determined, ranged from 70 percent to 90 percent except in years when the influenza season. Protective efficacy, so determined, ranged from 70 percent to 90 percent except in years when the circulating virus was antigenically different from the strains included in the vaccine (3).

While these evaluations were confirmed by more recent studies using similar endpoints and remain the “gold standard,” they were criticized even contemporaneously for not using virus isolation as an endpoint (4). The reason, partially theoretical, was based on the fact that vaccine would produce a rise in titer in the vaccinees only and that it might be more difficult to detect a further rise in titer among this group, a rise in titer in the vaccinees only and that it might be more difficult to detect a further rise in titer among this group, given infection, than in the unvaccinated. A concept of being near the “antibody ceiling” was even introduced as part of this concern, without much biologic justification. The major reason that virus isolation was not used, except to define the period of virus circulation, involved the logistics of specimen collection from all participants at the time of illness.

The problem in using virus isolation as an endpoint in large studies is that the procedure is more labor intensive and less comprehensive than using serology. Blood specimens for serologic studies can be easily collected from all participants before and after the season, while collecting specimens for virus isolation involves first recognizing an illness meeting the case definition and then contacting the case and carrying out the clinical procedure. There should not be a potential for selection bias in choosing who should be cultured in a blinded study, such as those carried out in the military, or, for that matter, in civilian populations when it may be more difficult to collect specimens in a planned fashion. Those who questioned using serology as the only endpoint proposed simply using differences in risk of all clinical respiratory illness meeting a case definition by vaccine status as a way of defining efficacy. Given the occurrence of acute respiratory illnesses of noninfluenzal etiology even in the influenza season, these analyses would always result in lower clinical efficacy estimates, biased toward no effect, although the exact degree of bias would vary.

Over the years, influenza vaccination became accepted and even recommended for certain segments of the civilian population. However, estimates of the efficacy of inactivated influenza vaccine again became a policy-related issue in the early 1990s when it was decided that only if there could be some demonstration of cost-effectiveness of the vaccine for the US Medicaid system would it be approved for coverage. The problem here was further complicated by the inability in the United States to conduct randomized, placebo-controlled trials in the target population, those 65 years of age or older, since, by then, influenza vaccination was an absolute recommendation for this age group. As a result, only observational studies could be conducted. These observational studies concentrated on hospitalization as an endpoint, since hospitalization was a great expense to the health care system and the investigation had an economic underpinning. The outcomes examined in those studies that used a case-control design were those discharge diagnoses corresponding to the categories classically defined as “pneumonia and influenza”
(5–7). Pneumonia and influenza outcomes have traditionally been used as a surrogate for hospitalizations or deaths that might be related to influenza. Other studies using cohort designs examined additional causes of hospitalization (8).

There was no attempt in these studies to identify influenza infection by laboratory tests, partially because of logistics but also because, by the time hospitalization occurs, influenza virus may be difficult to detect and secondary bacterial infection may have occurred. However, the cases included were limited to those hospitalized during the defined influenza season; in one study (6), a period of peak influenza transmission and a period of low influenza activity were compared. Any protection attributed to vaccination would be expected to be higher in the peak influenza season than in a period when transmission is at a low level. However, even in the peak influenza season, a portion of the hospitalizations of unknown size would not be caused by influenza, which would reduce the calculated effectiveness estimates. In these studies, the effectiveness of vaccination in preventing pneumonia and influenza hospitalizations was estimated at 30–57 percent. There was no way to adjust these estimates to determine true vaccine efficacy in preventing influenza-related hospitalizations only, because of difficulties in detecting whether influenza was the cause of an illness by the time hospitalization occurred. With the development of polymerase chain reaction techniques that may identify persistent nonreplicating virus, it might be possible today to overcome this problem of detecting recent infection (9). Without these techniques or any other way to identify the proportion of hospitalizations truly related to influenza, estimates of actual effectiveness resulting from these studies can involve only informed speculation.

**Efficacy studies of live-attenuated vaccines**

The live-attenuated vaccines represent a very different approach to preventing influenza. While somewhat newer than the inactivated vaccines, they have been under development for approximately 40 years (10). At first, the approach involved empiric “cold adaptation,” that is, using the technique first used for polio attenuation by simply growing the virus at reduced temperatures. The method of producing the vaccine viruses now involves taking advantage of the segmented genome of the virus, using reassortment techniques (11). Reassortment is required to update the virus as necessary, just as the inactivated vaccine is updated to reflect prevalent viruses in circulation. The reassortment results in all but the hemagglutinin and neuraminidase genes of the virus coming from a known cold-adapted, attenuated parent, one for type A viruses and one for type B, termed the “master strains.” The hemagglutinin and neuraminidase genes are derived from contemporary viruses, so the resulting vaccine, like the inactivated vaccines, contains two currently recommended type A viruses, A(H3N2) and (H1N1), and a type B virus.

Live-attenuated vaccines have been evaluated for many years, but the history of testing the current trivalent vaccine is much shorter. Most earlier studies did not evaluate clinical efficacy or, if they did, were not of sufficient size to give definitive results (12). It was learned that hemagglutination-inhibiting antibody response to the vaccine was not as good as with the inactivated vaccine, especially for A(H1N1) viruses, even though those persons with only modest antibody production appeared to be protected. This protection was thought, at least in part, to represent the unmeasured effects of cell-mediated immunity (13).

The definitive study of the vaccine in young children was conducted in the 1996–1997 influenza season (14). It used isolation-confirmed illness as the outcome and was directed at children 15–71 months of age. Positive cultures were thought to be the most reliable way to identify infection in these children. There was the old argument about the best way to identify infection in a vaccinated population, plus the fact that young children sometimes do not produce a strong antibody response to infection. The study was large, conducted among 1,602 children. The large size was fortunate and permitted evaluations of both influenza types, A(H3N2) and B, circulating that season, with efficacy estimates (and tight confidence intervals) of 95 percent (88–97 percent) and 91 percent (79–96 percent), respectively. No type A(H1N1) virus occurred during the study. In a second study year, A(H3N2) viruses circulated but of a strain that varied from the one included in the vaccine; results suggested an unexpected breadth of immunity as a result of vaccination (15). Again, endpoints were determined by virus isolation.

In contrast, a study conducted with this vaccine in adults did not use either virus isolation or rise in antibody titer to detect infection. This randomized blinded trial compared the live-attenuated vaccine with placebo in 4,561 healthy working adults (16). The idea was that if the clinical case definition used required signs of severe illness, such as fever, influenza would be positively selected and the results demonstrated would approximate results had virologic endpoints been used. This was not the case, possibly because the outbreak was of relatively small scale in much of the country and the A(H3N2) virus circulating that year was markedly different from the strain included in the vaccine (17). As a result, the sensitive but not specific primary outcome, febrile illness, was not significantly reduced among the vaccinated. However, febrile upper respiratory illness was reduced by 17 percent, and other outcomes were similarly reduced. These relatively low efficacy estimates were almost certainly also reduced because noninfluenza illnesses diluted the differences. Limiting analyses to the peak period of transmission raised the efficacy estimate for prevention of febrile upper respiratory illness to 24 percent. However, because no virologic determinations were conducted as part of the study, it is not possible to say what the result would have been had true influenza been identified. Had virus isolation been carried out in at least a subset of the participants, it might have been possible to take this into account by using methods recently described by Halloran and Longini (18).

**The Texas community studies**

The studies with the live-attenuated vaccine currently being conducted in Texas, reported in part in the accompanying article (19), were really designed to evaluate not direct
protection from use of the live vaccine in those vaccinated, but indirect protection of those not vaccinated, produced by reducing virus transmission. Such an approach was demonstrated to be effective in 1968 when schoolchildren in the community of Tecumseh, Michigan, were vaccinated to prevent transmission to the rest of the community (20). More recently, the live-attenuated influenza vaccine currently in use in the former Soviet Union was demonstrated to indirectly protect schoolchildren, that is, unvaccinated classroom contacts of vaccinated children (21).

Although the previously described studies of the trivalent live vaccine in children and adults were intended to assess direct protection of those vaccinated, the Texas study was aimed at determining the level of indirect protection produced. Three communities were chosen that were relatively similar. Each had a comparable type of managed health care system, and the endpoint chosen for comparison was medically attended illnesses. Vaccine was offered to all children in one of the communities, Temple-Belton, whether or not their families were members of the managed care organization. No vaccine was offered to children in the two other communities.

Occurrences of medically attended illness were compared between vaccinated and unvaccinated children in the intervention community, mainly to determine direct effectiveness, and between residents of the intervention community and the other two, mainly to quantify indirect protection, the goal of the study. As in past studies in which the primary outcome was not laboratory-confirmed influenza illness, it would be expected that the direct effectiveness estimates would be reduced by misclassification of the outcome, even in a period of defined influenza circulation. The amount of misclassification would not be known and would actually be greater than that found in the study involving adults, given the higher frequency of noninfluenzal illnesses in children. Issues of bias in a situation in which both participants and medical staff would know the vaccination status of participants need also be considered.

The accompanying article (19) illustrates a method to adjust for at least some of the misclassification that has occurred. Viral cultures, collected to help in timing the outbreak, were used to adjust estimates made by comparing risk of medically attended illness. This is a major contribution and shows how it is possible to take a study designed for another purpose and learn additional important information. The authors also examined the degree to which their adjusted estimates were influenced by the fact that vaccine might have reduced the severity of influenza illness in the vaccinated group, thus reducing the likelihood that these subjects would be sampled. Because cultures were being collected for surveillance purposes, the unvaccinated might have been oversampled, not necessarily because of greater illness severity but possibly because the medical staff might have thought that the unvaccinated would be more likely to be influenza positive.

The sensitivity analysis is of particular interest, and it illustrates that the method would not be unduly influenced by the way those cultures were selected. The large increase in the effectiveness estimates using culture results for adjustment is not surprising given the efficacy estimates determined in the randomized clinical trials. It is basically a result of the much higher frequency of positive cultures in the unvaccinated group as a proportion of collected specimens. However, uncertainties are introduced by the need to use non-randomly-collected specimens. In addition, because of the relatively small number of specimens collected, the confidence intervals around the estimates are wide.

Given these wide intervals, some of which include zero, it seems rash to try to overinterpret the data beyond that which can be demonstrated convincingly. In particular, this statement applies to those vaccinated in 1999, the year before the outbreak, but not in 2000, the year of the outbreak. Both the age-specific data for those vaccinated in 1999 (presented in table 3 of the paper by Halloran et al. (19)) with the types combined and the data presented for type A (H1N1) and type B separately are not significant, while those for 2000 are. Thus, caution should be exercised when interpreting the point estimates for the possible 2-year protection produced by the vaccine, especially since their interpretation needs to be coupled with speculation about the relatively minor effects of changes between the viruses in the vaccines and those circulating. Long-term protection may be present, but this analysis has not proven it. What has been demonstrated is that the vaccine appears to protect against type A (H1N1) illness, the subtype that did not circulate during previous trials. These results can be added to those demonstrated in an experimental challenge study designed to determine whether the live-attenuated vaccine truly protects against A (H1N1) virus (13). Thus, we can safely conclude that the vaccine is highly effective not only against type A (H3N2) and type B but also type A (H1N1) in children vaccinated in the autumn before the influenza season.

The contribution of the paper (19) is not simply the determination of this efficacy but the methods used. These methods, while useful, cannot correct for the various biases, many of which cannot be quantified, which may be present in a nonblinded study of this type. Only a blinded randomized trial can do that. When a study is designed to answer one question and is used to answer another, it is not usually possible to get definitive results. This analysis does not do the impossible but does produce interesting information supporting results in other studies. The method may even be more useful when it is included in advance as part of the study protocol. Virus culture could be obtained from a fixed proportion of those vaccinated and not vaccinated, whatever the study design. Thus, use of culture would extend the interpretation of results of trials in which this outcome is used only in a fraction of participants and improve the accuracy of what might otherwise be a nonvirologic endpoint.

REFERENCES