Case Series Analysis of Adverse Reactions to Vaccines: A Comparative Evaluation

C. P. Farrington, 1 J. Nash, 2 and E. Miller 3

A modified cohort method has been proposed for estimating the relative incidence of rare adverse reactions after vaccination. The method requires only a sample of the cases, thus avoiding the need for following large population cohorts or selecting controls. This case series method has statistical power equivalent to that of the full cohort method when the risk periods after vaccination are short and vaccine coverage is high. The method also eliminates confounding by variables associated with both the outcome and avoidance of vaccination. In this paper, the cohort, case-control, and case series methods are reviewed, and their underlying assumptions and performances are compared. Theoretical results are illustrated using data on febrile convulsions after measles-mumps-rubella vaccination in the United Kingdom. Am J Epidemiol 1996;143:1165-73.

case-control studies; cohort studies; epidemiologic methods; statistics; vaccination

Vaccine safety with respect to rare adverse events is an issue of substantial public and epidemiologic interest. However, while safety with respect to common reactions may be established in clinical trials, these studies do not usually possess sufficient power for testing hypotheses of association with rare adverse events. Epidemiologic studies are therefore required to investigate suspected or hypothesized associations once the vaccine is in widespread use.

Such studies present several methodological difficulties. Many potential sources of bias have been identified, such as differential ascertainment of cases in recently vaccinated and unvaccinated individuals and differential vaccination rates in individuals at higher or lower risk (1). Cohort studies must be very large to achieve sufficient power, and may therefore be impractical, though it is sometimes possible to undertake retrospective cohort studies in large data sets assembled for other purposes (2). Case-control studies require much smaller overall sample sizes, but are open to bias in the selection of controls. Both cohort and case-control studies may suffer from confounding by variables related both to avoidance of vaccination and to the outcome of interest. Thus, for instance, Fine and Chen (1) list parental education, ethnic group, age of the mother, maternal smoking, birth weight, evolving neurologic disorders, and conditions predisposing to seizures as factors related both to vaccination and to sudden infant death syndrome or encephalopathy. Few studies could realistically aspire to control for all of these factors.

An alternative study design using only data on cases has been proposed for investigating acute adverse events (3). This design combines the power and simplicity of the cohort method and the economy of the case-control method, while reducing confounding. The method is similar in spirit to the case-crossover design proposed by Maclure (4) for studying the effect of transient exposures on myocardial infarction and to that described by Feldman (5) for assessing drug reactions. It differs from these methods in that it is derived from the same statistical model as the cohort study design, from which it inherits many properties, including the facility for handling several dependent exposures and controlling for age effects, both of which are essential in the present context. It also avoids the need to specify the prior probabilities of exposure (6). The case series method has been used to estimate the relative incidence of febrile convulsions after diphtheria-tetanus-pertussis (DTP) and measles-mumps-rubella (MMR) vaccination in the United Kingdom (7).
This paper compares the cohort, case-control, and case series methods, with particular emphasis on their assumptions, relative efficiencies, and handling of confounding. The performance of the three methods is contrasted in a study of febrile convulsions after MMR vaccine in the United Kingdom, using a subset of the data reported by Farrington et al. (7).

THE THREE METHODS

The adverse reactions considered in this paper are nonspecific acute events, and do not include possible long term effects of vaccination. An association with vaccination is manifested by a temporal clustering or trough of events shortly after vaccination. The relation between vaccination and the outcome of interest may be conveniently described by the proportional incidence model:

\[ \lambda(t) = \rho \lambda_0(t), \]

where \( \lambda(t) \) and \( \lambda_0(t) \) denote, respectively, the incidence at time \( t \) in the presence and absence of recent vaccination and \( \rho \) is the relative incidence. The vaccine effect is assumed to operate in one or several relatively short risk periods after vaccination. Thus, for instance, a study of DTP vaccine and febrile convulsions or encephalopathy used the risk periods 0–3, 4–7, 8–14, and 15–29 days after any dose of DTP (8). Risk periods may correspond to different levels of risk or to different reaction mechanisms, based, for instance, on evidence gathered from passive reporting systems. Prior definition of risk periods is essential in order to avoid the temptations and pitfalls of data-dependent choices.

The cohort method

In the cohort method, a subsample of the population is followed over time. When vaccine coverage is high, the vaccinated and unvaccinated populations are likely to differ in important respects, and a direct comparison of adverse event rates in vaccinated and unvaccinated groups may be biased. Instead, rates are compared in different periods with respect to vaccination. An observation period is defined during which individuals are followed. The incidence rates for each risk period are estimated using person-time denominators and are compared with the incidence rate for the control period. This is defined as the person-time within the observation period that is not included in a risk period. Adjustments for age are made by subdividing each period into age groups in which the incidence is roughly constant. The analysis may be stratified by other covariates, and the adjusted relative incidences are estimated by log-linear modeling. The method has been used extensively in retrospective cohort studies (8–11).

The cohort approach may be used both for unique events, such as sudden infant death syndrome, and for potentially recurrent events, such as febrile convulsions. For unique events, the person-time at risk in cases ends with the event, whereas for recurrent events, the entire observation period is included in the person-time denominators. Inclusion of second and subsequent episodes is valid provided that events are independent. If this is not to be the case, the analysis can be restricted to first episodes. For unique events, the analysis can also use survival methods, with age and vaccine effects entering the model as time-dependent variables.

The case-control method

The case-control method provides a practical way of estimating relative incidence when events are rare. All cases or a random sample of cases are assembled over a given time period. For each case, one or several matched controls are selected, usually among individuals who have not experienced the event of interest up to the end of this period. Matching is necessary in order to determine exposure in controls. For each case-control set, a reference age is defined as the age of the case when the event occurred. Exposure is defined as vaccination within a specified time interval prior to the reference age; several exposure categories may be used. The exposure period corresponds to the risk period after vaccination in the cohort method. Thus, for example, in the National Childhood Encephalopathy Study of DTP vaccine and serious acute neurologic events (12), five exposure periods were defined: 0–3, 4–7, 8–14, 15–28, and ≥29 days prior to the reference age. To control for the variability in vaccination coverage over calendar time, it is usual to match cases and controls closely on date of birth. Further matching according to potentially confounding factors, such as geographic area, ethnic group, and socioeconomic status, is advisable but may be difficult to achieve. Recurrent events may be handled by defining exposure as vaccination within any of the exposure intervals determined by the ages at which events occurred in the case.

Conditional logistic regression is used to estimate odds ratios, which are indistinguishable from the relative incidences when the background incidence is low and the exposure periods are short. When close matching on date of birth is not required, a modified case-control method known as the case-cohort design can also be used (13, 14). With this method, stratified subsamples of the cohort are used as control groups for the cases in each stratum, thus increasing efficiency.

Am J Epidemiol Vol. 143, No. 11, 1996
However, this analysis is more complex, since allowance must be made for the nonindependence of vaccination in different time intervals for each control.

The case series method

The cohort method is based on comparisons of incidence rates for person-time aggregated both across and within individuals. The case series method removes the contribution of comparisons between individuals, focusing attention on event rates in different periods within each individual's observation time. Individuals who experience no events therefore no longer contribute any information about the association between vaccination and the outcome, and can be ignored without introducing any bias. Individuals who do experience one or more events, on the other hand, contribute information on the risk period and age group in which the events occurred. Combining this information across cases results in an estimate of relative incidence. Technically, the case series model is derived from a cohort model by conditioning the analysis on the total number of events observed for each individual (3).

The method is applied as follows. An observation period of calendar time is defined within which all events or a random sample of events are selected. As for the cohort method, person-time of observation for each individual is assigned to age and risk categories, together with the number of events occurring in each category. Unlike the cohort method, the various categories are not aggregated across individuals. The analysis is performed by multinomial modeling, conditional on the observed number of events and the vaccination history for each case. This may be achieved by log-linear modeling with the ungrouped data, exactly as for the cohort method but including an additional factor with a level for each individual to keep the total number of events experienced by each individual fixed at its observed value.

Implementation of the case series method

The case series method is easily implemented, since only a sample of cases arising during a specified calendar time period of observation is required, along with the vaccination histories of the sample cases up to the end of the observation period. Only individuals with known vaccination histories are included in the analysis. Also required is the date of birth of each case and a personal identifier with which to link multiple events in the same individual. Data on covariates such as sex, location, or socioeconomic status may be collected to investigate the impact, if any, of those factors on the presumed vaccine effect.

In common with the other methods discussed in this paper, the case ascertainment mechanism must be independent of vaccination history. Thus, reporting systems for adverse events that are temporally related to vaccination, such as the Vaccine Adverse Event Reporting System (15), are inappropriate for this purpose. In the United Kingdom, the method has been applied using databases of hospital admissions (7) and laboratory diagnoses (3, 16). Similar databases are available in the United States, and may be augmented by data linkage (17). Thus, the Vaccine Safety Data-link Project coordinated by the Centers for Disease Control and Prevention aims to link adverse events and vaccination records for the purpose of evaluating pediatric vaccine safety, using a variety of analytical tools, including cohort and case series methodology (Robert Chen, Centers for Disease Control and Prevention, unpublished manuscript).

COMPARISON OF THE METHODS

Assumptions

The cohort model is based on the assumption that events arise in an age-dependent Poisson process. In its simplest form, the background incidence is assumed to be the same for all individuals. The case series model is derived from a cohort model in which the background incidence is allowed to differ between individuals. In this sense, the basic assumptions are weaker for the case series method than for the cohort method. However, for unique events, the case series method requires the additional assumption that the cumulative incidence of events in the population over the observation period is low (3). The three statistical models are shown in the Appendix. The proportional incidence assumption is essential to the case series and case-control methods, but may be relaxed for the cohort method.

If, in the case of recurrent events, second and subsequent events cannot be regarded as independent, the analysis should be restricted to first events. The case series analysis may then be applied as for unique events.

Adjustments to all three methods are required if vaccination is not independent of previous adverse events—for instance, if the occurrence of an adverse event constitutes a contraindication for vaccination. The observation period for the cohort and case series methods should then begin with age at vaccination, while for case-control studies, only the first event should be used to define the reference age.

Heterogeneity of the relative incidence between different population groups—between males and females, for instance—may be analyzed with all three
methods by means of the interaction between the relevant covariate and the vaccine effect. However, heterogeneity of the background incidence cannot be estimated with the case series method, or indeed with the case-control method if cases and controls are matched on the covariate of interest.

**Power**

For a given number of events, the case series method is only slightly less powerful than the cohort method, provided that the proportion of individuals who receive vaccine during the observation period is high. This is established theoretically in a simple case by means of the asymptotic relative efficiencies of the log relative incidence for the case series method, previously derived (3), and of the log odds ratio for the case-control method, relative to a cohort analysis based on the same number of events. The efficiency of an estimate is defined as the inverse of its variance. (The formulae are given in the Appendix.) The relative efficiency of the case series method is close to 1 when the risk period is short in comparison with the observation period, or when the population vaccine coverage in the observation period is high. In particular, if vaccine coverage is 100 percent, the cohort and case series methods have the same efficiency. Figure 1 shows how the relative efficiencies depend on the values of the parameters; relative incidences greater than 1 indicate an increased risk after vaccination. The relative efficiency of both the case series method and the case-control method declines with increasing relative incidence. However, when vaccine coverage is high, the decline in relative efficiency for the case series method is small. Thus, for the purpose of estimating relative incidences, little information is lost if noncases are ignored. Indeed, in situations where incidence varies between individuals, the case series method removes this extraneous variability and can therefore be more powerful than the cohort method. In comparison, the case-control method is considerably less efficient, although its power may be improved by using several controls per case, or by case-cohort sampling (14).

It follows that there is little to be gained by assembling data for the entire cohort, since the case series

---

**FIGURE 1.** Efficiency of the case series method (top three curves in each section) and the case-control method (bottom three curves), relative to the cohort method, in detecting rare adverse vaccine reactions, by relative incidence. Proportion of the population vaccinated in the observation period: ---, 90%; . . . . ., 50%; -- --, 10%. Proportion of the observation period at risk: left section, 5%; right section, 20%.
analysis based on the same number of events provides broadly the same power. A sample size formula for the case series method is given in the Appendix. This allows calculation of the required number of events for a given power and significance level. The design parameters to be supplied are the relative incidence that needs to be detected, the proportion of the population vaccinated in the observation period, and the ratio of the risk period to the average observation period.

Confounding

Fine and Chen (1) have highlighted the important issue of bias in studies of vaccine safety. Potential confounding from factors associated with the outcome and with avoidance of vaccination is eliminated in the case series design, since unvaccinated individuals contribute no information on the relative incidence associated with vaccination, though they do contribute information about relative incidence associated with age. Nevertheless, the case series method remains susceptible to bias if vaccination is timed to minimize the risk of an adverse event. Investigating such effects is not easy whatever the study design. One possibility is to investigate interactions between the relative incidence and factors indicating a predisposition to the adverse event. Assembling the required data is likely to be easier in case series than in cohort or case-control studies.

Absolute incidence

The cohort approach, which is population-based, produces estimates of absolute incidence, whereas the case series and case-control methods do not. However, provided that the cases constitute a census of those arising out of a defined population, rough estimates can be obtained, orders of magnitude for the absolute risk being sufficient for most practical purposes. This contrasts with the need for carefully calculated confidence intervals and p values in order to assess associations between outcomes and vaccination. In many situations, however, it may not be possible to relate the cases to a defined population—for instance, if cases are ascertained from admissions to a hospital that accepts patients referred from other areas. In such a situation, it would be difficult to design a retrospective cohort study or to identify correctly the population from which controls should be sampled for a case-control study. The case series approach, however, sidesteps this difficulty entirely, since it requires no explicit definition of the cohort from which the cases arise.

EXAMPLE: METHODS

The relative efficiency of the three methods will be illustrated here in a real data set obtained retrospectively reconstructing a population cohort. The cases for this study comprise a subset of those from a larger study reported on by Farrington et al. (7). This larger study was designed specifically for the case series method. However, in one area, the denominator data for a cohort study were also collected, thus providing an opportunity to compare methods. The present analysis was undertaken solely as a comparative exercise; the reader is referred to the paper by Farrington et al. (7) for a discussion of the epidemiologic issues.

Construction of the cohort

All admissions for febrile convulsion occurring in children aged 366–730 days from January 1, 1991, to February 28, 1993, were extracted from the administrative systems of the two district general hospitals in the former South East Kent Regional Health Authority. Readmissions occurring within 72 hours with the same diagnosis were counted as the same episode.

All vaccination records of children living in South East Kent who were born between January 1, 1989, and February 28, 1992, and who received at least one dose of DTP, poliomyelitis, or MMR vaccine were obtained from the South East Kent Regional Health Authority Child Health Vaccine Database. This database includes all children residing in the area and is used to schedule vaccination appointments, mail reminders to parents, and record vaccinations. All MMR vaccines used in South East Kent contain the Jeryl Lynn strain of mumps virus.

Vaccination and admission records were matched using a combination of computerized and manual procedures based on each child's name, date of birth, sex, and postal code. The matching procedure is described elsewhere (18). This yielded a retrospectively assembled cohort for which we had information on date of birth, sex, dates of vaccination, and dates of admission to hospital for febrile convulsion. For each individual, the observation period thus began on January 1, 1991, or at age 366 days, whichever was latest, and ended on February 28, 1993, or at age 730 days, whichever was earliest. Repeat admissions for febrile convulsion within this period were included.

Cohort analysis

In the cohort analysis, two postvaccination risk periods were considered: 6–11 days and 15–35 days after MMR vaccine. These are the risk periods potentially associated, respectively, with the measles and
mumps components of the vaccine (7, 16). All other days were included in the control period. Age was grouped into six broadly equal intervals. The total numbers of convulsions and total person-time in each age, sex, and risk category were calculated. Relative incidence was estimated using log-linear modeling with factors for age, sex, and risk period.

**Case-control analysis**

In the case-control analysis, children aged 366–730 days who had been admitted to hospital for febrile convulsions were listed in random order. For each case, in turn, a single control child was randomly selected from all children of the same sex in the population cohort with a birth date within 30 days of that of the case, and not matched to previous cases on the list. The reference ages for the case-control pair were defined as the age or ages of the case child upon admission for febrile convulsion. Two exposure levels were defined: MMR vaccination within 6–11 days or within 15–35 days prior to any of the reference ages.

Conditional logistic regression was used to obtain the adjusted odds ratios for each exposure level. Since the events considered were rare, the odds ratios were indistinguishable from the relative incidences. To allow for the sampling variability among the controls, we repeated the entire control selection procedure and analysis 100 times using a computer program, and the median estimates were used for comparison with the other methods.

**Case series analysis**

The case series method used only data on children admitted to hospital for febrile convulsions between the ages of 366 and 730 days. The same risk periods and age groups were used as for the cohort method. For each child, the total number of convulsions and person-time in each risk and age category were calculated. Relative incidence was estimated using multinomial regression that was conditional on the number of convulsion episodes experienced by each child.

**EXAMPLE: RESULTS**

Of the 113 admissions for febrile convulsion among children aged 366–730 days that were extracted from the hospital administration databases, 105 (93 percent) were successfully matched to immunization records. The eight nonmatching admissions were from individuals who did not reside in the area. Two readmissions within 72 hours were discounted, and the remaining 103 admissions from 91 patients were included in the analysis. Six children experienced two episodes of convulsions, and three experienced three episodes. The denominator obtained from the immunization database included 11,815 individuals, of whom 11,431 (97 percent) received MMR vaccine prior to February 28, 1993, the end of the observation period. Eighty-seven of the 91 cases (96 percent) received MMR vaccine prior to that date. There were a total of 2,929,534 person-days of observation, including 44,994 in the 6- to 11-day risk period and 158,874 in the 15- to 35-day risk period after MMR vaccination.

Three of the 103 convulsion episodes occurred 6–11 days after MMR vaccination, and six occurred 15–35 days after vaccination. Table 1 shows the relative incidences obtained using the cohort method, based on the cohort of 11,815 children, and the case series method, based on the 91 cases, together with median estimates obtained from 100 case-control studies with one matched control for each of the 91 cases. The 95 percent confidence intervals for each risk period included 1 for all three methods of estimation. In 27 of the 100 case-control studies, no controls were exposed in the 6- to 11-day period, thus leading to unbounded relative incidences. In these cases, the exact univariate lower 95 percent confidence limits were computed. Figure 2 shows the distribution of the 100 estimates and their 95 percent confidence intervals.

**DISCUSSION**

The relative incidence estimates and confidence intervals obtained using the cohort and case series methods were practically identical, as predicted by the

<table>
<thead>
<tr>
<th>Risk period after MMR vaccination</th>
<th>Cohort method (estimated RI)</th>
<th>Case-control method (median OR)</th>
<th>Case series method (estimated RI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6–11 days</td>
<td>1.99 (0.61–6.48)</td>
<td>2.77 (0.28–27)</td>
<td>2.02 (0.63–6.45)</td>
</tr>
<tr>
<td>15–35 days</td>
<td>1.10 (0.47–2.59)</td>
<td>1.25 (0.34–4.65)</td>
<td>1.06 (0.46–2.54)</td>
</tr>
</tbody>
</table>

* MMR, measles-mumps-rubella; RI, relative incidence; OR, odds ratio; CI, confidence interval.
FIGURE 2. Estimated log odds ratios (○) from 100 replications of the control selection procedure, in increasing order, by risk period after measles-mumps-rubella vaccination (6–11 days (top) and 15–35 days (bottom)). Vertical lines, 95% confidence intervals.

theory. Indeed, the confidence intervals were slightly narrower for the case series method, due to the variation in background incidence between individuals. This source of variability is eliminated by the case series method. The median estimates produced by the case-control method were similar to those given by the other two methods, but the confidence intervals were substantially wider, reflecting the lower power of the method for a given number of cases. The relative incidence estimates obtained from these data were similar to those from the larger study (7). The larger study found a significantly raised incidence of febrile convulsions 6–11 days after MMR vaccination and 15–35 days after receipt of MMR vaccine containing the Urabe strain of mumps virus, but no evidence of an increased incidence 15–35 days after receipt of vaccine containing the Jeryl Lynn strain (7), as used in South East Kent.

In conclusion, this example confirms the theoretical finding that in highly vaccinated populations, the case series method has statistical power equivalent to that of the cohort method and outperforms the case-control method. This finding has practical implications for the study of vaccine reactions, a field that has hitherto been restricted by the sample size requirements of cohort designs and the potential for biased selection of controls in case-control designs. The case series method also avoids the problem in retrospective cohort
studies of defining the population from which the cases arose. In addition, self-matching removes confounding due to variables related to both the outcome and avoidance of vaccination.

The cohort study remains the “ideal” design for the study of adverse reactions to vaccines, and should be used whenever feasible. However, for studies of rare adverse events or for routine surveillance purposes, large-scale cohort studies may be costly, impractical, or prone to confounding. In such circumstances, the case series method provides a powerful and practical alternative to cohort and case-control studies.

REFERENCES

APPENDIX
Let individuals be indexed by i, age groups by j, and risk or exposure periods by k. Let \( \lambda_{ijk} \), \( e_{ijk} \), and \( n_{ijk} \) denote, respectively, the incidence, time at risk, and number of events experienced by individual i in age group j and risk period k. Suppose the incidence is parameterized using the simple log-linear model:

\[
\ln(\lambda_{ijk}) = \phi_i + \alpha_j + \beta_k.
\]

For the cohort (co) model, \( \varphi_i = x_i^T \gamma \) for fixed covariates \( x_i \), and the Poisson log-likelihood kernel (which is equal to the log-likelihood up to an additive constant) is

\[
\text{lik}_c = -\sum_{ijk} \exp(x_i^T \gamma + \alpha_j + \beta_k) e_{ijk} + \sum_{ijk} n_{ijk}(x_i^T \gamma + \alpha_j + \beta_k).
\]

The case series (cs) model is derived from a cohort model with the \( \varphi_i \) unrestricted by conditioning on the \( n_{i} \), thus giving the multinomial log-likelihood kernel:

\[
\text{lik}_c = \sum_{i} \sum_{ijk} n_{ijk} \log \left( \frac{\exp(\alpha_j + \beta_k) e_{ijk}}{\exp(\sum_k \beta_k x_{ik})} \right).
\]

See Farrington (3) for details. Now consider a case-control (cc) study with one control per case, matched on age and possibly other covariates. The log-likelihood kernel is then

\[
\text{lik}_c = \sum_{i} \log \left( \frac{\exp(\sum_k \beta_k x_{ik})}{\exp(\sum_k \beta_k x_{ik}) + 1} \right),
\]

where \( \beta_k \) now represents the log odds ratio of exposure in the kth exposure category. The sum is over all case-control pairs, with \( x_{ik} = 1 \) if the ith case was and the corresponding control was not vaccinated in the kth exposure period, \( x_{ik} = -1 \) if the converse occurred, and \( x_{ik} = 0 \) if case and control were both vaccinated or not in the kth exposure period.

For the purpose of deriving relative efficiency measures, consider a simplified situation in which there are no age effects and a single risk period, and all individuals are observed for the same period of time and share a common background incidence. Suppose also that a proportion \( \nu \) are vaccinated during this time, and that the ratio of the risk period to the observation
period is \( r \). Let \( \rho \) denote the relative incidence \( e^\beta \) and \( n \) the number of events. Let \( \text{var}_{co} \), \( \text{var}_{cs} \), and \( \text{var}_{cc} \) denote, respectively, the variance of the estimate of \( \beta \) obtained from the cohort method, the case series method, and the matched case-control method with a unique control per case. Inverting the information matrix obtained as minus the second derivative of the log-likelihood shows that, when the total number of events \( n \) is large and the incidence is low,

\[
\text{var}_{co} = n^{-1} \left[ 1 + \nu r (\rho - 1) \right] \left[ (1 - \nu r)^{-1} + (\nu r)^{-1} \right];
\]

\[
\text{var}_{cs} = n^{-1} \left[ 1 + \nu r (\rho - 1) \right] \left[ (1 - r)^{-1} + (\rho r)^{-1} \right] \nu^{-1};
\]

\[
\text{var}_{cc} = n^{-1} (1 + \rho) \left[ (1 - \nu r)^{-1} + (\nu r)^{-1} \right],
\]

and hence the relative efficiencies (RE) of the case series method and case-control method with respect to the cohort method, for a given number of cases, are

\[
\text{RE}_{cs} = \frac{1 + \nu r \rho / (1 - \nu r)}{1 + \rho \rho / (1 - r)};
\]

\[
\text{RE}_{cc} = \frac{1 + \nu r (\rho - 1) / (1 + \rho)}{1 + \nu r (\rho - 1) / (1 + \rho)}.
\]

If all individuals are vaccinated during the observation period, the cohort and case series methods have the same asymptotic efficiency.

The formula for \( \text{var}_{cs} \) may be used to derive a sample size formula. The number of events required to detect a relative incidence \( \rho \) with power \( 100(1 - \beta)\% \) at the \( 100\alpha\% \) significance level, given a population vaccination coverage \( \nu \) in the observation period and a risk period \( r \) as a proportion of the total observation period, is

\[
n = \frac{1}{\nu r (1 - r) (\rho - 1)^2} \times \left\{ Z_{\alpha/2} + Z_{\beta} \sqrt{[1 + \nu r (\rho - 1)] [1 + r (\rho - 1) / \rho]} \right\}^2,
\]

where \( Z_{\alpha/2} \) and \( Z_{\beta} \) are the corresponding centiles of the standard normal distribution.