Table 8.2. Vaccine efficacy under leaky and all-or-none models of vaccine action as measured in case-control studies. f_1: sampling fraction for cases, f_2: sampling fraction for controls, VE_k: vaccine efficacy calculated with "not yet cases" as controls, VE_r: vaccine efficacy calculated with "total population" as controls (adapted from Smith, Rodrigues, and Fine 1984).

		Cases		Not yet cases		VE_k	Total population		VE_r
Leaky model									
Year 3	Vaccinated	44	f_1	905	f_2	73%	1000	f_2	64%
	Unvaccinated	121	f_1	670	f_2		1000	f_2	
Year 6	Vaccinated	38	f_1	779	f_2	73%	1000	f_2	64%
	Unvaccinated	67	f_1	368	f_2		1000	f_2	
All-or-none model									
Year 3	Vaccinated	31	f_1	918	f_2	81%	1000	f_2	74%
	Unvaccinated	121	f_1	670	f_2		1000	f_2	
Year 6	Vaccinated	17	f_1	842	f_2	89%	1000	f_2	75%
	Unvaccinated	67	f_1	368	f_2		1000	f_2	

population that is vaccinated. If any two of the variables are known, the other can be estimated. Let PCV be the proportion of cases that are vaccinated and PPV be the proportion of the population that is vaccinated. Then vaccine efficacy can be estimated using the screening method by the following relation:

$$\mathrm{VE}_S = 1 - \frac{PCV(1-PPV)}{(1-PCV)PPV}. \tag{8.6}$$

Farrington (1993) discusses the screening method and its relation to case-control studies. An estimate of vaccine coverage in the population provides an estimate of the proportion vaccinated. Not all cases need to be ascertained, but the cases ascertained should be a random sample. Bias and precision of the method are considered. A method for computing a confidence interval allowing for extra variability and a method to determine sample size are provided. The screening method offers a simple, rapid and inexpensive surveillance tool to get approximate estimates of vaccine effectiveness. It is called a screening method because it can be used to suggest when further more accurate evaluation of a vaccine in the field might be needed.

8.2 Validation Sets for Outcomes

8.2.1 Validation sets for outcomes in vaccine studies

In vaccine field studies, often a nonspecific case definition rather than a more specific confirmatory diagnosis is used as the outcome. As seen in Chapter 6,

estimates of VE_S will be lower when based on less specific case definitions, particularly when the diagnosis of the disease of interest is not biologically-confirmed. Sometimes it would be prohibitively expensive or invasive to confirm each suspected case in a study biologically. However, if the biological confirmation can be done in a small random sample of the suspected cases, then this information can be used to estimate the expected number of cases of the true disease of interest among the suspected cases. The added uncertainty from not confirming all of the cases is taken into account with the statistical method, leading to a larger variance and wider confidence interval than would be obtained if all suspected cases were biologically confirmed.

In essence, this is a particular case of a missing data problem. The outcome of interest may be measured on some of the study participants in a subset called a validation sample, while the less specific outcome is measured on all participants. Then the result of the biological outcome is missing in those nonspecific cases who were not tested. In this situation, statistical missing data methods are available to use the outcomes of interest in the validation sample to correct the low estimates based on the nonspecific case definition alone. Statistical methods that use specific measures in small samples of the study subjects to correct bias when using nonspecific measures in the main study are used in other epidemiologic fields. Using validation sets for exposure to infection to improve joint estimation of VE_S and VE_I could also be useful (Golm et al, 1998, 1999). In the next sections, the use of validation sets and the gain they can provide is illustrated using data from an ongoing observational study of trivalent live-attenuated influenza vaccine in Central Texas.

8.2.2 Influenza vaccine field study in central Texas

Halloran et al (2003) evaluated the protection of a trivalent, live attenuated, influenza virus vaccine (LAIV-T) against influenza during the influenza season of 2000-01. They used surveillance cultures taken from a sample of the study participants to obtain more accurate estimates of protective efficacy against influenza than those obtained using the nonspecfic, clinical case definition.

A field study of LAIV-T was conducted in Temple-Belton, Texas, and surrounding areas during the 2000-01 influenza season. The field study was part of a larger community-based, non-randomized, open-label field study conducted from 1998-2001 in Temple-Belton, Texas, as well as two other communities to evaluate the indirect effectiveness of LAIV-T vaccination of healthy children (Gaglani t et al. 2004; Piedra et al 2005a, Piedra et al 2005b). Temple-Belton was the intervention community. At that time, the Temple-Belton area had approximately 19,700 children from 18 months through 18 years of age. In Temple-Belton, eligible healthy children and adolescents aged 18 months through 18 years were offered LAIV-T vaccine through the Scott & White (S & W) Clinics from 1998-2001. The analysis using validation set methods includes children who were S & W Health Plan (SWHP) members, and is

Table 8.3. Study data for influenza epidemic season 2000-2001 (from Halloran *et al.* 2003).

Age (years)	Vaccine Status	Children	MAARI cases	MAARI proportion	MAARI cases cultured	Number positive cultures	Fraction cultures positive	Fraction cultured
1.5-4	LAIV-T	537	389	0.72	16	0	0	0.041
	None	1844	1665	0.90	86	24	0.28	0.052
5-9	LAIV-T	807	316	0.39	17	2	0.12	0.054
	None	2232	1156	0.52	118	53	0.45	0.102
10-18	LAIV-T	937	219	0.23	19	3	0.16	0.087
	None	5249	1421	0.27	123	56	0.46	0.087
Total	LAIV-T	2281	924	0.41	52	5	0.10	0.056
	None	9325	4242	0.45	327	133	0.41	0.077

concerned with the LAIV-T vaccinations administered in the influenza season 2000-01. Children received a single dose of LAIV-T each year that they enrolled. Age-eligible members of the SWHP on January 7, 2001 were considered for inclusion in the analysis. Age at time of vaccination was used for those who received LAIV-T during the 2000-01 season. For those who never received LAIV-T, the first day of enrollment for the 2000-01 season, November 6, 2001, was used to compute age.

The primary clinical outcome was a non-specific case definition called medically-attended acute respiratory infection (MAARI), which included all ICD-9-CM diagnoses codes (Codes 381-383, 460-487) for upper and lower respiratory tract infections, otitis media and sinusitis. Any individual presenting with history of fever and any respiratory illness at S & W Clinics was eligible to have a throat swab (or nasal wash in young infants) for influenza virus culture. The decision to obtain specimens was made irrespective of whether a patient had received LAIV-T. The specific case definition is culture-confirmed influenza. Table 8.3 contains the number of children, the number of MAARIs, the number of cultures done, and the number of cultures positive for each group. The overall fraction of MAARI cases sampled was a little higher in the unvaccinated than in the vaccinated groups for those vaccinated in 2000 ($p = 0.03$). As expected, the proportion of cultures that were positive was consistently higher in the unvaccinated than in the vaccinated groups.

The risk of developing MAARI was compared in the children receiving LAIV-T with those children who had never received LAIV-T. The protective effectiveness of LAIV-T against MAARI was estimated as $\mathrm{VE}_{S,CI,a} = 1 - \mathrm{RR}$, where RR is the relative risk of MAARI in vaccinated children compared to unvaccinated children. The "a" stands for auxiliary outcome. Age-adjusted estimates were obtained using sample size weighted averages. Confidence intervals were based on the assumption of a normal approximation of the logarithm of the ratio of two independent binomial random variables (Katz 1978).

8.2.3 Analysis using surveillance samples

Estimates of the protective efficacy of LAIV-T against influenza using the surveillance samples, $\mathrm{VE}_{S,CI,v}$, were obtained using the mean score method for auxiliary outcomes (Pepe et al 1994), an estimating equations approach for handling missing data. The "v" stands for validation sample. The method estimates the score contribution for main study members with only auxiliary outcome data from the mean of the score contributions of a sample of study subjects with the same observed covariate and auxiliary outcome values on whom the specific outcome has been measured. In this analysis, the clinical outcome MAARI was the nonspecific, auxiliary outcome, while the actual influenza status was the specific outcome of interest. The confidence intervals take into account the uncertainty due to culturing only a sample of the MAARI cases.

The variable Y = outcome of interest (influenza status), A = auxiliary outcome (MAARI, yes or no), X = set of covariates (vaccination, age group), $P_\beta(Y|X)$ = binomial probability model, β = parameters to estimate in the probability model, S_β = score function, and $V, \overline{V}$ = in the validation set or not. The estimating equation is

$$\sum_{i \in V} S_\beta(Y_i|X_i) + \sum_{j \in \overline{V}} \hat{E}\{S_\beta(Y|X_j)|A_j, X_j\} = 0.$$

An unbiased estimator for a person who had no culture done is:

$$\hat{E}\{S_\beta(Y|X_j)|A_j, X_j\} = \sum_{i \in V(A_j, X_j)} S_\beta(Y_i|X_i)/n^V(A_j, X_j).$$

The variance was estimated on the adjusted log relative risk using the mean score and multivariate delta methods (Pepe et al 1994; Agresti 1990; Chu et al 2003). With just one covariate, as in this situation, the model is saturated, so the mean score method is equivalent to the semiparametric efficient method (Rotnitzky and Robins 1995).

Let $R = 0, 1$ denote whether the influenza status is known or missing. Children with positive MAARI who were sampled for influenza cultures have $R = 1$. Those with positive MAARI who were not sampled for influenza culture have $R = 0$, and the influenza status is missing. The analysis assumed that all children with negative MAARI were also negative for influenza disease, and thus, $R = 1$. The mean score method produces valid estimates if the data are missing at random (MAR) (Pepe et al 1994) in the sense of Little and Rubin (2002). In our example, MAR means that Y and R are conditionally independent given A and X, denoted $(Y \perp R \mid A, X)$. If Y and R are conditionally independent given (A, X), then $[Y|X, A, R = 1] = [Y|X, A]$, where the brackets denote a probability density distribution (Clayton et al 1998).

A continuity correction of 0.5 was added to the number of cultured samples and the number positive in the age group 1.5-4 years in the analysis of vacci-

Table 8.4. Epidemic year 2000-2001: Vaccine effectiveness ($VE_{S,CI,a}$) against MAARI and efficacy ($VE_{S,CI,v}$) against combined influenza A (H1N1) and B taking missing influenza status into account.

Age (years)	$VE_{S,CI,a}$ MAARI	(95% CI)	$VE_{S,CI,v}$ influenza	(95% CI)
1.5-4	0.20	(0.14,0.25)	0.91	$(-0.34, 0.99)$
5-9	0.25	(0.15,0.34)	0.80	(0.26,0.95)
10-18	0.14	(0.01,0.26)	0.70	(0.13,0.90)
Total	0.18	(0.11,0.24)	0.79	(0.51,0.91)

nation in both years because there were no positive cultures in the vaccinated group.

The protective efficacy estimates against influenza taking missing influenza status into account are much higher than the estimates of the protective effects of LAIV-T against MAARI (Table 8.4). The overall vaccine effectiveness estimate based on the nonspecific case definition was 0.18 (95% CI: 0.11,0.24). The overall efficacy estimates incorporating the surveillance cultures using the mean score method was 0.79 (95% CI:0.51,0.91). In this situation, using the surveillance cultures as a validation set resulted in a four-fold increase in estimates, much closer to the efficacy estimate of 0.93 (95% CI: 0.88,0.97) obtained in a double blind randomized controlled trial (Belshe *et al.* 1998, Section 6.4.3). Although the point estimates are higher, the confidence intervals are wider due to the uncertainty resulting from not culturing all of the MAARI cases.

Table 8.4 contains the overall estimate obtained by pooling the data and avoiding the continuity correction. The age-adjusted $VE_{S,CI,v}$ obtained using sample size weighted averages, the continuity correction in the youngest age group, and the delta method for the variance estimate was $VE_{S,CI,v} = 0.77$, [95% CI 0.48,0.90], similar to that in Table 8.4.

In this study, selection of children with MAARI for influenza cultures was not done randomly. Physicians might tend to choose MAARI cases that they believe to be influenza for culturing. If influenza disease is more moderate in the vaccinated group, then oversampling in the unvaccinated group might occur based on the influenza status, which is not measured on everybody. In this case, the MAR assumption is violated and the estimate assuming MAR could be biased.

If physicians know vaccination status, they might oversample either the unvaccinated or the vaccinated children. They might tend to believe that vaccinated children would not have influenza, and therefore oversample the unvaccinated children. However, oversampling due to knowledge of the vaccination status alone would not bias the estimate, since the estimation procedure stratifies on the vaccination status of the child. The data are missing at random in this case (Little and Rubin 2002). In fact, in future studies, it

would be desirable to oversample the vaccinated, non-specific cases for culturing. Oversampling in the vaccinated group would help avoid having zero positive cultures in the vaccinated groups.

The consistently higher proportion of cultures being positive in the unvaccinated groups could be partly due to vaccinated cases of influenza being less likely to be culture positive than unvaccinated cases. However, this would produce exactly the same bias that would be obtained if all of the MAARI cases had been cultured as in many randomized, double-blinded vaccine trials.

Future vaccine field studies that utilize validation samples could be intentionally designed so that the specific outcome would be missing at random within any given observed stratum of the study subjects. The sample size needed in the validation sample to correct the bias from using the non-specific outcome is not necessarily large. In this case the overall sampling fraction was well below 10 percent. However, with a highly efficacious vaccine, one might need to oversample in the vaccinated groups to avoid structural zeros.

8.3 Sensitivity analysis for validation sets

8.3.1 Sensitivity analysis for selection bias

The analysis in Section 8.2 relies on the non-identifiable assumption that the outcome of interest is missing at random (MAR) (Little and Rubin, 2002). If the outcome is not MAR, then effect estimates could be subject to selection bias. Rotnitzky *et al.* (1998, 2001), Scharfstein *et al.* (1999), and Robins *et al.* (2000) developed a frequentist selection model that displays the sensitivity analysis over a plausible range of selection parameters. Scharfstein *et al.* (2003) developed a Bayesian approach that allows the formal incorporation of prior beliefs about the degree of the selection bias on the odds ratio scale to obtain the full posterior distribution, a single summary of the sensitivity analysis. Scharfstein et al (2006) extended this work to the relative risk parametrization of selection bias, discrete covariates, and dependence of the priors for the relative risk parameters across treatment groups. They re-analyzed the data from the Texas influenza study (Section 8.2.2) with the methods. They relied on an influenza expert to provide informative priors for the Bayesian analysis.

8.3.2 Sensitivity analysis in the vaccine study

In the vaccine field study, let n be the total number of participants, and n_0 and n_1 the number of vaccinated and non-vaccinated participants, respectively. Let Z denote the vaccination indicator, taking on the value 1 if a participant is vaccinated and 0 if not vaccinated. Let $A(0)$ and $A(1)$ denote the indicator of MAARI (1: yes; 0: no), for a participant if she had been, possibly contrary to fact, unvaccinated or vaccinated, respectively. The observed MAARI outcome $A = A(Z)$ is observed for every participant. Let $Y(0)$ and $Y(1)$ denote

influenza status (1: positive; 0: negative) for a participant if she had been, possibly contrary to fact, unvaccinated and vaccinated, respectively. Only one of these outcomes can be potentially observed. In this study, influenza status is biologically confirmed by a culture. In the validation sub-study, a possibly non-random sample of the participants are biologically confirmed, so that influenza status, $Y = Y(Z)$ is known for a subset of the participants. Let R be the validation indicator, where $R = 1$ if sampled for validation and $R = 0$, otherwise. Sampling for validation only occurs for those with $A = 1$. Let X denote age category (0: 1.5-4 years; 1: 5-9 years; 2: 10-18 years) measured at the time of study entry.

The observed data for an individual are $O = (Z, X, A, R, Y : A = R = 1)$. We assume we observe n i.i.d. copies, $\boldsymbol{\mathcal{O}} = \{O_i : i = 1, \ldots, n\}$. Throughout, probabilities P, indexed by subgroup subscripts indicate restriction to the associated subpopulation. For example, for events A and B, $P_{z,x}[A] = P[A|Z = z, Z = x]$ and $P_{z,x}[A|B] = P[A|B, Z = z, X = x]$.

Vaccine Efficacy

The scientific goal is to use the observed data to estimate the causal effect of vaccination on the outcome Y, within age levels as well as overall. Specifically, the goal is to estimate age-specific vaccine efficacy

$$VE_{S,CI,x} = 1 - \frac{P_x[Y(1) = 1]}{P_x[Y(0) = 1]} \tag{8.7}$$

and overall vaccine efficacy

$$VE_{S,CI} = 1 - \frac{\sum_{x=0}^{2} P_x[Y(1) = 1]P[X = x]}{\sum_{x=0}^{2} P_x[Y(0) = 1]P[X = x]}. \tag{8.8}$$

To identify $VE_{S,CI,x}$, it is sufficient to identify $P_x[Y(z) = 1]$ for $z = 0, 1$. For $VE_{S,CI}$, we must identify $P_x[Y(z) = 1]$ for all z and x, and the marginal distribution of X. While the marginal distribution is identified from the observed data without additional assumptions, the conditional probabilites $P_x[Y(z) = 1]$ will require non-identifiable assumptions.

Two structural assumptions facilitate identification of $P_x[Y(z) = 1]$. The first is that Z is independent of $\{A(0), A(1), Y(0), Y(1)\}$ given X. This assumption states that vaccination status is independent of the potential outcomes $\{A(0), A(1), Y(0), Y(1)\}$, given age (X). That is, within levels of age, vaccination is randomized. Although the Texas influenza study was not randomized, the expert had no information to conclude that one group differed substantially from another.

The second assumption is that $A(z) = 0$ implies $Y(z) = 0$. This assumption is needed because the study design called for passive case ascertainment. As a result, no participants who do not appear in the clinic are cultured for confirmation of influenza infection. The above assumption states that if

a participant, under vaccination status z, does not have MAARI, then she does not have medically-attended influenza. The interest is in efficacy against medically-attended, culture-confirmed influenza, not influenza infection.

Identification of $P_x[Y(z) = 1]$

With these assumptions, we can write

$$P_x[Y(z) = 1] = P_{z,x}[Y(z) = 1] \tag{8.9}$$
$$= P_{z,x}[Y = 1] \tag{8.10}$$
$$= \sum_{r=0}^{1} P_{z,x}[Y = 1|A = 1, R = r]P_{z,x}[A = 1, R = r]. \tag{8.11}$$

Equation (8.9) follows from randomization within levels of X; (8.10) uses the fact that $Y = Y(Z)$; and (8.11) follows from an application of the law of conditional probability and the second assumption above. For all z, x, r, $P_{z,x}[Y = 1|A = 1, R = 1]$ and $P_{z,x}[A = 1, R = r]$, are identifiable but $P_{z,x}[Y = 1|A = 1, R = 0]$ are not. Thus, identification of $P_x[Y(z) = 1]$ will require identification of these latter probabilities.

The most common assumption used to identify these probabilities is that of missing at random (MAR) (Little and Rubin 2002). MAR states that R is independent of Y given (Z, A, X). This implies that, for all z, x, $P_{z,x}[Y = 1|A = 1, R = 0] = P_{z,x}[Y = 1|A = 1, R = 1]$. As a result, $P_x[Y(z) = 1]$ becomes identifiable. Since the assumption of MAR is untestable and was considered questionable by the scientific expert, it is useful to perform a sensitivity analysis to outcomes that are missing not at random (MNAR).

8.3.3 Frequentist Sensitivity Analysis

Scharfstein *et al.* (1999, 2003) and Robins *et al.* (2000) introduced a sensitivity analysis methodology in which a class of models (including MAR) are posited, each yielding identification of $P_{z,x}[Y = 1|A = 1, R = 0]$. They recommended that inferences about the estimands of interest be presented over a range of posited models, considered plausible by subject-matter experts. In particular, they assumed a pattern-mixture model. Scharfstein et al (2006) showed that the pattern-mixture model could be re-written as a selection model. In the selection model, for subjects with $Z = z, X = x$ and MAARI, the selection bias parameter $\alpha_{z,x}$ is interpreted as the log odds ratio of being unvalidated for diseased vs. undiseased subjects. So, $\alpha_{z,x}$ positive or negative indicates that diseased subjects have lower or higher odds of being validated, respectively.

When eliciting plausible ranges for $\alpha_{z,x}$, the expert found it easier to think about selection bias on a relative risk as opposed to an odds ratio scale. Specifically, he felt more comfortable expressing opinions about the relative risk of being validated given that a MAARI participant has influenza, compared with

having another influenza-like illness. As a result, the selection models were reformulated in terms of the relative risk selection bias parameters

$$\beta_{z,x} = \frac{P_{z,x}[R=1|A=1,Y=1]}{P_{z,x}[R=1|A=1,Y=0]}. \tag{8.12}$$

Then specification of $\beta_{z,x}$ leads to identification of $P_x[Y(z)=1]$ via the following formula:

$$\begin{aligned} P_x[Y(z)=1] &= \frac{P_{z,x}[Y=1|A=1,R=1]P_{z,x}[A=1]}{\beta_{z,x}P_{z,x}[Y=0|A=1,R=1]+P_{z,x}[Y=1|A=1,R=1]} \\ &= \frac{P[Z=z,X=x,A=1,R=1,Y=1]P[Z=z,X=x,A=1]/P[Z=z,X=x]}{\beta_{z,x}P[Z=z,X=x,A=1,R=1,Y=0]+P[Z=z,X=x,A=1,R=1,Y=1]} \end{aligned} \tag{8.13}$$

The frequentist non-parametric estimator of $P_x[Y(z)=1]$ can be found by replacing the probabilities P in (8.13) by their empiricals $\tilde{P}$. Plugging the estimates $\widehat{P}_x[Y(z)=1]$ into equations (8.7) and (8.8) yields the estimates $\widehat{VE}_{S,CI,x}$ and $\widehat{VE}_{S,CI}$. The right hand side of estimated equation (8.13) reduces to the results with the mean score method when $\beta_{z,x}=1$, for all z and x. Supplementary web material for Scharfstein et al (2006) give the derivation of the large sample based confidence intervals for $VE_{S,CI,x}$ and $VE_{S,CI}$. The frequentist sensitivity analysis proceeds by varying the $\beta_{z,x}$ over plausible ranges.

Figure 8.1 shows the estimated probabilities (and 95% confidence intervals) for influenza in the vaccinated and unvaccinated groups, for each age stratum, as a function of $\beta_{z,x}$. Figure 8.2 shows the point estimates and lower 95% confidence bounds for the age-group-specific efficacy. The selection bias parameters were varied over the 90% ranges elicited from the expert. Within these ranges and within each age group, the vaccine efficacy estimates based on the validation sets are higher than the point estimates based on the non-specific definition, which were 0.2, 0.25, and 0.14 for the age groups 1.5-4, 5-9, and 10-18, respectively. The lower confidence bounds indicate the degree of variability. A key drawback of the frequentist sensitivity analysis methodology is that it is not feasible to present parsimoniously the overall results. This is a motivation for the Bayesian sensitivity analysis.

8.3.4 Bayesian inference

To simplify notation, we let $\boldsymbol{\beta}_z = (\beta_{z,0},\beta_{z,1},\beta_{z,2})'$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_0',\boldsymbol{\beta}_1')$, $\boldsymbol{\eta}_z = (\eta_{z,0},\eta_{z,1},\eta_{z,2})'$, $\boldsymbol{\eta} = (\boldsymbol{\eta}_0',\boldsymbol{\eta}_1')'$, $p_{z,x} = P_{z,x}[Y=1|A=1]$, $\boldsymbol{p}_z = (p_{z,0},p_{z,1},p_{z,2})'$, $\boldsymbol{p} = (\boldsymbol{p}_0',\boldsymbol{p}_1')$, $\phi_{z,x,a} = P_z[A=a,X=x]$, $\boldsymbol{\phi}_z = (\phi_{z,0,0},\phi_{z,0,1},\phi_{z,1,0},\phi_{z,1,1},\phi_{z,2,0},\phi_{z,2,1})'$, and $\boldsymbol{\phi} = (\boldsymbol{\phi}_0',\boldsymbol{\phi}_1')'$. With this notation,

$$P_x[Y(z)=1] = P_{z,x}[Y=1] = p_{z,x}\frac{\phi_{z,x,1}}{\sum_{a=0}^{1}\phi_{z,x,a}},$$

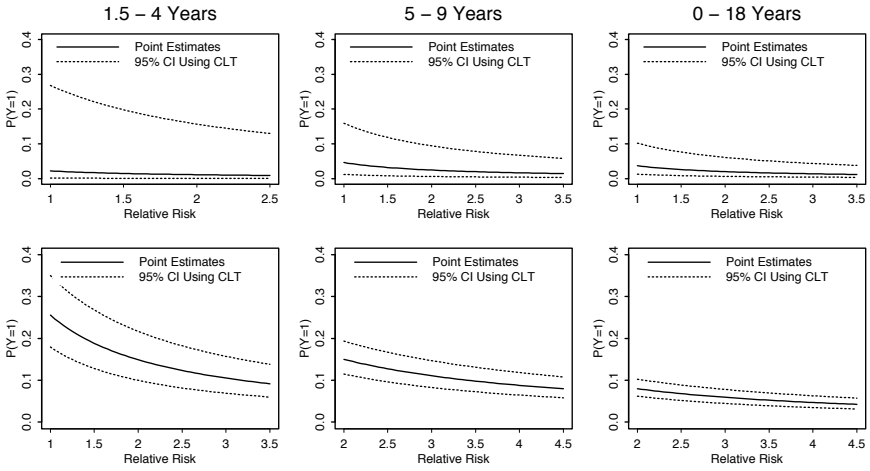

Fig. 8.1. Frequentist sensitivity analysis of $P_{z,x}[Y = 1]$. Shown are the estimated probabilities (and 95% confidence intervals) for influenza in the vaccinated and unvaccinated groups, for each age stratum, as a function of the relative risk selection bias parameter $\beta_{z,x}$, varied over the 90% ranges elicited from the expert.

and

$$VE_{S,CI,x} = 1 - \frac{p_{1,x}}{p_{0,x}} \frac{\phi_{1,x,1}}{\phi_{0,x,1}} \frac{\sum_{a=0}^{1} \phi_{0,x,a}}{\sum_{a=0}^{1} \phi_{1,x,a}}, \tag{8.14}$$

$$VE_{S,CI} = 1 - \frac{\sum_{x=0}^{2} \left\{ p_{1,x}\phi_{1,x,1} \left(\sum_{z=0}^{1} \sum_{a=0}^{1} \phi_{z,x,a} \right) / \left(\sum_{a=0}^{1} \phi_{1,x,a} \right) \right\}}{\sum_{x=0}^{2} \left\{ p_{0,x}\phi_{0,x,1} \left(\sum_{z=0}^{1} \sum_{a=0}^{1} \phi_{z,x,a} \right) / \left(\sum_{a=0}^{1} \phi_{0,x,a} \right) \right\}} \tag{8.15}$$

In the prior specification for $\boldsymbol{\beta}$, Scharfstein et al (2006) provide two options: (1) Bayesian analogue of the frequentist sensitivity analysis and (2) fully Bayesian analysis. For option (1), point-mass priors were specified on $\boldsymbol{\beta}$ and the posterior distributions of the estimands of interest were estimated over a range of $\boldsymbol{\beta}$. This approach is comparable to the frequentist sensitivity analysis described above, but does not rely on large sample approximations. For option (2), a non-degenerate prior distribution on $\boldsymbol{\beta}$ was specified as elicited from a subject-matter expert. This approach has the advantage of providing a single summary of the posterior inference about the estimands, which naturally incorporates the uncertainty due to selection bias.

Details of the specification of the joint prior distribution on $(\boldsymbol{\beta}', \boldsymbol{\eta}', \boldsymbol{p}', \boldsymbol{\phi}')'$ are in Scharfstein et al (2006). To sample from the posterior, they constructed a Gibbs sampling algorithm with data augmentation (Tanner and Wong, 1987) and slice sampling (Damien *et al.* 1999). All programs were written in R. The MCMC algorithm was run for 500,000 iterations with 100,000 discarded as burn-in.

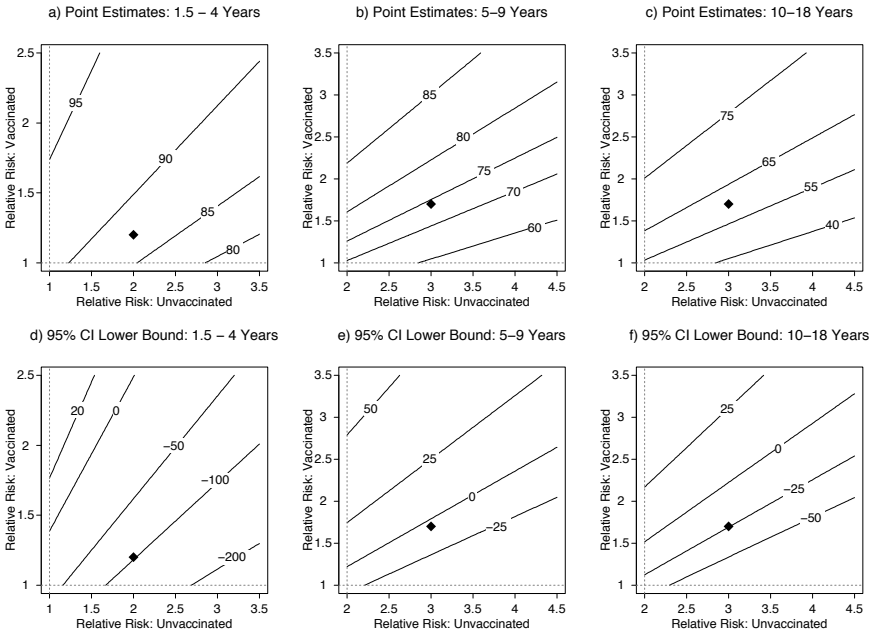

Fig. 8.2. Frequentist sensitivity analysis of point estimates and lower 95% confidence bounds for the age-group-specific vaccine efficacy as a function of the relative risk selection bias parameters $\beta_{1,x}$ (vaccinated) and $\beta_{0,x}$ (unvaccinated) varied over the 90% ranges elicited from the expert. Black diamonds indicate the results at the best guess of the expert. Black lines with numbers indicate the contours.

Informative priors

For Bayesian inference, informative priors are specified for the selection bias relative risk parameters, $\boldsymbol{\beta}$, by age group and vaccination status. An influenza expert was asked the following question: *If a physician were doing surveillance cultures during an influenza season, what is the probability that he would select the children who actually had true influenza over the children who just had nonspecific respiratory symptoms to culture?* He responded that this was very hard to answer. One "would be more likely to be correct in the unvaccinated," because unvaccinated children presenting with true influenza would have more typical, severe disease than the vaccinated children. One would be "less likely to be correct in young children under five years," because children under five years experience many other severe respiratory diseases that could be mistaken for influenza, while older children are already immune to such diseases. He added that the degree of selection bias would also "depend on the rules for collection, for example, a certain number per week or with specific symptoms."

Another influenza expert had similar views. He provided his best guess for each of the univariate relative risk selection bias parameters $\beta_{z,x}$ defined in

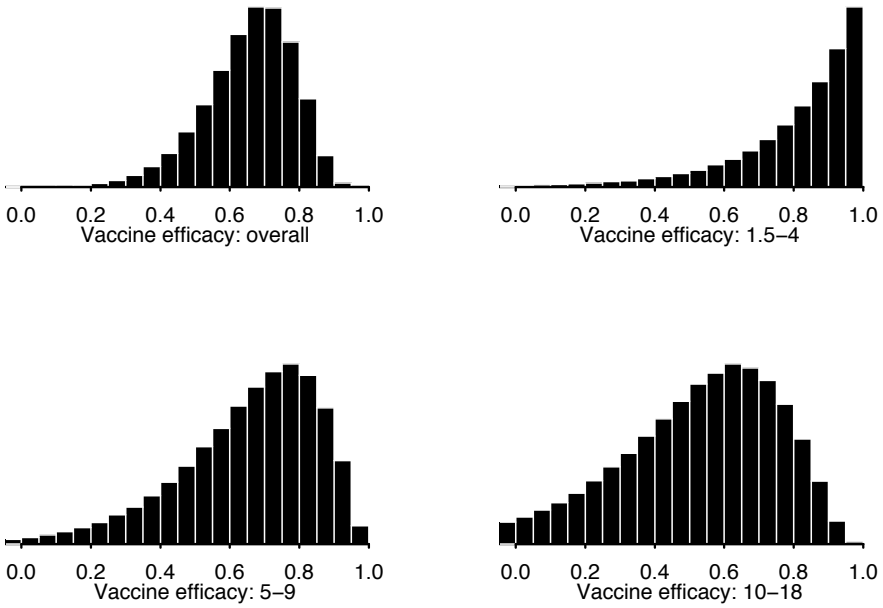

Fig. 8.3. Posterior distributions of the overall vaccine efficacy and by age-group using the informative prior distributions.

(8.12) and his belief about the interval that would likely contain 90% of the prior distribution for each $\beta_{z,x}$. He also provided the degree of correlation of the selection bias by age group and vaccine status. The expert did not have different prior beliefs about the selection bias in the 5-9 and 10-18 year age groups. Thus the prior distributions for these two age groups are presented as one group in Table 8.5.

For the analysis, the elicitations were plugged into a multivariate Normal prior on the $\log \boldsymbol{\beta}$ scale as follows. The elicited best guesses for each $\beta_{z,x}$ and 90% range were transformed to the $\log \beta_{z,x}$ scale. Using `qnorm` in Splus, the univariate Normal(mean, sd) distributions that best approximated the elicited priors on the log scale were found. In the unvaccinated 5-18 year olds, the elicitation was quite consistent with a Normal distribution. In the other three groups, adjustments were necessary as shown in Table 8.5. The expert felt comfortable with the changes in elicited and proposed priors in light of the uncertainty about the selection bias. Table 8.5 contains the log transformed elicited priors, the slightly adjusted univariate Normal priors with their 90% interval on the $\log \beta_{z,x}$ scale, and the corresponding best guess and ranges on the relative risk, $\beta_{z,x}$, scale. The corresponding univariate distributions used for the $\log \beta_{z,x}$ for the unvaccinated $1.5 - 4$ year old and older children are

Table 8.5. Best guess and 90% range for the informative prior distributions on the selection bias parameter $\boldsymbol{\beta}$ and log $\boldsymbol{\beta}$

Age		$\boldsymbol{\beta}$ (relative risk) scale				log $\boldsymbol{\beta}$ scale			
group		Unvaccinated		Vaccinated		Unvaccinated		Vaccinated	
(years)		best guess	90% range	best guess	90% range	best guess	90% range	best guess	90% range
$1.5-4$	elicited	2.00	1.00,3.50	1.20	1.00,2.50	0.69	0.00, 1.25	0.18	0.00,0.92
	used	2.00	1.00,4.00	1.20	0.58,2.50	0.69	0.00, 1.39	0.18	−0.55,0.92
$5-18$	elicited	3.00	2.00,4.50	1.60	1.00,3.50	1.10	0.69,1.50	0.47	0.00, 1.25
	used	3.00	2.00,4.50	1.70	1.00,2.89	1.10	0.69,1.50	0.53	0.00, 1.06

Normal(0.69, 0.43) and Normal(1.10, 0.25), and for the vaccinated $1.5-4$ year olds and older children are Normal(0.18,0.57) and Normal(0.53, 0.32).

The expert believed that the correlation in selection bias among the strata would be quite high, even as high as 0.90. The corresponding covariance matrices for $\pi(\boldsymbol{\beta})$ were constructed from the marginal univariate Normal distributions and the correlations.

Figure 8.3 shows the Bayesian posterior distribution of the age-group-specific efficacy and overall efficacy using the informative prior distributions from Table 8.5, assuming a correlation of 0.9. The mode is 1.00 in the age group 1.5-4 years, since there are no positive cultures in the vaccinated group in that age-group. The results assuming a zero correlation were nearly identical (not shown).

Table 8.6 compares the summaries of the Bayesian posterior distributions and of the frequentist estimates and 95% confidence intervals. The assumption of MAR results in an overestimate of the vaccine efficacy compared with the selection bias relative risk assumptions elicited from the expert. The frequentist result with β fixed at 1 (MAR) is identical to the result in Table 8.4. The Bayesian result with $\pi(\beta)$ fixed at 1 is the same as in Chu and Halloran (2004). The Bayesian posterior means are somewhat lower than the frequentist estimates. In general, the Bayesian credible intervals are tighter than the frequentist confidence intervals. This is especially true for the age-group 1.5 to 4 years, where the validation sample is small and there are no positive cultures.

8.4 Validation sets with time-to-event data

8.4.1 Texas influenza vaccine field study 2003–2004

The Texas field study (Section 8.2.2) continued in the fall of 2003. Children were not vaccinated in the 2002-2003 season. In the meantime the vaccine

Table 8.6. Results of Bayesian and frequentist sensitivity analyses using surveilance cultures assuming NMAR and MAR. For the Bayesian analyses, the posterior means (95% credible intervals) for vaccine efficacy are reported, for the frequentist analyses, the point estimates (95% confidence intervals). The estimates using just the nonspecific MAARI case definition are included for comparison.

	Age Group			
Analysis	1.5-4 years	5-9 years	10-18 years	Overall
Bayesian:				
informative $\pi(\boldsymbol{\beta})$	0.80 (0.23,1.00)	0.65 (0.13,0.93)	0.51 (−0.12,0.88)	0.65 (0.35,0.86)
$\pi(\boldsymbol{\beta})$ fixed at best guess	0.77 (0.11,0.99)	0.63 (0.10,0.93)	0.50 (−0.12, 0.86)	0.64 (0.32,0.85)
$\pi(\boldsymbol{\beta})$ fixed at 1 (MAR)	0.84 (0.41,0.90)	0.73 (0.40,0.94)	0.64 (0.26,0.89)	0.73 (0.53,0.88)
Frequentist:				
β fixed at best guess	0.88 (−0.97,0.99)	0.74 (−0.05,0.88)	0.61 (−0.25,0.88)	0.73 (0.34,0.89)
β fixed at 1 (MAR)	0.91 (−0.34,0.99)	0.80 (0.26,0.95)	0.70 (0.13,0.90)	0.79 (0.52,0.91)
MAARI alone:	0.20 (0.14,0.25)	0.25 (0.15,0.34)	0.14 (0.01,0.26)	0.18 (0.11,0.24)

was licensed but not approved for children under 5 years old. In the 2003-2004 season, healthy children and adolescents aged 5 – 18 years were eligible to enroll and were offered LAIV-T vaccination at public schools, the Temple Mall, churches, and Scott & White (S & W) Clinics. In the 2003-04 influenza season, the predominant circulating influenza A (H3N2) virus in the United States was similar antigenically to A/Fujian/411/2002 (H3N2), a drift variant of A/Panama/2007/99 (H3N2), the vaccine strain. Comparison of the vaccinated with the unvaccinated children within Temple-Belton allows prospective evaluation of the direct protective effects of LAIV-T against the drift variant during the 2003-2004 influenza season. Children who were contraindicated to receive LAIV-T, such as history of asthma, were offered trivalent inactivated vaccine (TIV). Thus, there were three main vaccinated groups of interest: 1) those receiving LAIV-T in 2003, whether or not they had received LAIV-T before, 2) those having received LAIV-T in the seasons 1998-99 through 2001-02, but not in 2002-03 or in the fall of 2003, 3) those receiving TIV in the fall of 2003. The distributions of chronic obstructive pulmonary diseases and other similar potential confounders were similar in the LAIV-T, the previously vaccinated, and the unvaccinated groups. The TIV group had a much higher percentage of COPD and other diseases than the other groups, so that comparison of the TIV group with the unvaccinated group is not valid. Age-eligible members of the SWHP on October 10, 2003 were considered for inclusion in the analysis. The final inclusion was restricted to children living within zip codes in the Temple-Belton area. The definition of a case of medically-attended acute respiratory illness (MAARI) is the same as in Section 8.2.2.

Central Texas influenza surveillance was performed as previously described (Gagliani et al 2004; Piedra et al 2005; Halloran et al 2003). Some of the children who had surveillance cultures done were in the SWHP, and others

were not. Those in the SWHP were included in the SWHP administrative database. The non-SWHP children were not in the SWHP database, though their culture results, age, and vaccination status were available. The primary influenza season was defined as the weeks with the most intense influenza activity accounting for 80-85% of all positive influenza cultures (Nichol et al 1999; Piedra et al 2005). The primary influenza season occurred during the 10-week period from October 10 – December 20, 2003. We considered MAARI cases and cultures within the 10-week period of interest. The influenza season started early in Texas, so vaccination occurred during the influenza season. This time-dependent analysis allowed children to change their vaccination status during the study period.

8.4.2 Time-to-event analysis

The effectiveness of protection against MAARI and the efficacy of protection against culture-confirmed influenza were computed using the basic equation $\mathrm{VE}_{S,IR} = 1 - RR$, where RR is the ratio of the number of MAARI (estimated influenza) cases/ child-days in the vaccinated compared to the unvaccinated group. In this section, we denote it simply as VE. Our main interest was in the efficacy of LAIV-T, but estimates were also obtained for TIV and previously vaccinated in 1998-2001 (PREV), all three being compared to the unvaccinated group. Age-group specific values were computed for the two age groups 5-9 years and 10-18 years. Overall efficacy was computed by weighting the contributions of the age groups by the combined number of child-days at risk in the vaccinated and unvaccinated groups in each age group.

A child who began the season as either unvaccinated or previously vaccinated could be switched to the LAIV-T group or the TIV group once they got vaccinated in 2003. To take the changing vaccine status into account and to integrate the surveillance cultures into the analysis, we grouped the data by week over the ten week period. If vaccination occurred before the day of MAARI, the child was counted as a vaccinated MAARI case. Otherwise, the child was counted as a previously vaccinated or unvaccinated MAARI case. We assumed that multiple visits in a week were not independent. Only the first MAARI case in the week was included if a child had more than one MAARI presentation in that week. Vaccine effectiveness against MAARI was denoted as VE^a.

To estimate the efficacy against culture-confirmed influenza illness, the expected number of influenza cases in each week for each age and vaccine group was estimated by multiplying the proportion of positive surveillance cultures in each age and vaccine group by the number of MAARI cases in that group (Halloran and Longini 2001). Children who had positive cultures were considered to be no longer at risk for influenza and did not contribute further child-days at risk for the rest of the ten week period.

The data are grouped within one-week time intervals τ, $(t_{\tau-1}, t_\tau]$, $\tau = 1, \ldots, T$, $T = 10$. Let k, $k = 1, \ldots, K$, indicate the relevant strata, in our case

Table 8.7. MAARI events, child-days at risk and rate per 1,000 child-days at risk by age group and vaccine status.

Age (years)	Vaccination status	MAARI Events	Child-days at Risk	Rate/1,000 Child-days at risk
5-9				
	LAIV-T	105	35,886	2.93
	TIV	80	10,598	7.55
	PREV	143	26,902	5.32
	UNVAC	261	61,522	4.24
10-18				
	LAIV-T	117	42,991	2.72
	TIV	82	13,741	5.97
	PREV	273	71,424	3.82
	UNVAC	641	179,828	3.56
Combined				
	LAIV-T	222	78,883	2.81
	TIV	162	24,383	6.64
	PREV	416	98,297	4.23
	UNVAC	902	241,331	3.74
Totals				
5-9		589	134,908	4.37
10-18		1113	307,984	3.61
Combined		1702	442,896	3.84

age groups, and $K = 2$. Let $n_{\nu\tau}$, $\nu = 0, 1$, be the number of participants in the unvaccinated and vaccinated group at risk of influenza at the beginning of each time interval, with $n_{k\nu\tau}$, $\nu = 0, 1$, $k = 1, \ldots, K$, the corresponding number in each stratum. For each stratum k, $k = 1, \ldots, K$, and vaccine status ν, $\nu = 0, 1$, let the number of MAARI cases ascertained in each time interval be $w_{k\nu\tau}$, the number of surveillance cultures be $r_{k\nu\tau}$, and the number of positive cultures be $c_{k\nu\tau}$. For each τ, estimate the proportion $\rho_{k\nu\tau}$ of true influenza cases among the MAARI cases in each age and vaccine group by $\hat{\rho}_{k\nu\tau} = c_{k\nu\tau}/r_{k\nu\tau}$. Multiply $w_{k\nu\tau}$ by $\{\hat{\rho}_{k\nu\tau}\}$ to obtain an estimate of the number of influenza cases in each interval. Summing over the weekly estimates of the number of true influenza cases, the total expected number of influenza cases in each age and vaccine group during the study is estimated. The outcome of interest, the result of a culture, is assumed to be missing at random (MAR) (Little and Rubin 2002).

To compute child-days at risk, everyone is assumed to be at risk for influenza at the beginning of the study period. For each time interval τ, the child-days at risk in each stratum, $d_{k\nu\tau}$, was computed as $7 \times (n_{k\nu\tau} - 0.5\hat{\rho}_{k\nu\tau}w_{k\nu\tau})$, $\nu = 0, 1$, $k = 1, \ldots, K$. That is, the expected number of influenza cases times half the time interval was subtracted from the number at risk at the beginning of the interval to adjust the child-days at risk. Children who had positive cultures were considered to be no longer at risk for influenza

and did not contribute further child-days at risk for the rest of the ten week period. The incidence rate of true influenza in each vaccine group and was estimated based on the above, and from that, the stratum specific vaccine efficacy, $\mathrm{VE}_{k,v}$, based on the validation set as

$$\widehat{\mathrm{VE}}_{S,IR,k,v} = 1 - \frac{[\sum_{\tau=1}^{T} \hat{\rho}_{k1\tau} w_{k1\tau}]/[\sum_{\tau=1}^{T} d_{k1\tau}]}{[\sum_{\tau=1}^{T} \hat{\rho}_{k0\tau} w_{k0\tau}]/[\sum_{\tau=1}^{T} d_{k0\tau}]} \ . \tag{8.16}$$

Overall $\mathrm{VE}_{S,IR,v}$ is computed by weighting the contributions of the age groups by the combined number of child-days at risk in the vaccinated and unvaccinated groups in each age group.

Two different estimates using the surveillance cultures can be computed. The first, denoted VE^{in}, uses just the surveillance cultures from the SWHP members in the database. The second, denoted VE^{ex}, uses the surveillances cultures from both the SWHP members and the non-SWHP children. The surveillance cultures from the SWHP members are called the internal validation set because we also have the MAARI data on these children, while the others are the external validation set. Confidence intervals were obtained using 2000 bootstrap samples (Efron and Tibshirani 1993). To get confidence intervals for VE^{ex}, the external culture data were bootstrapped separately, then each external bootstrap dataset was added to the corresponding bootstrap dataset from the main SWHP data set to get the bootstrap estimate of the proportion of cultures that were positive. Then VE^{ex} was computed for the bootstrap data set. In using the external cultures, the assumption was made that the population producing the non-SWHP cultures was similar to the population producing the SWHP data. When spread over a ten week period, the culture data were too sparse to obtain a separate weekly estimate of $\rho_{k\nu\tau}$ for use in equation (8.16). So the overall estimated proportion positive in Table 8.8 in each group was used as the estimate for the proportion positive in equation (8.16).

A total of 6,403 age-eligible children in the SWHP database living in the zip codes in the Temple-Belton are included in the analysis, of whom 1,706 received LAIV-T and 548 received TIV in 2003 before the end of the study period. Of the remaining children, 983 had been previously vaccinated in 1998-2001 and 3,166 had never been vaccinated by the end of the study period. About four weeks into the period, by November 8, 2003, 50% of the vaccinees had been vaccinated with either LAIV-T or TIV.

Table 8.7 contains the number of MAARI events, child-days at risk, and rate per 1,000 by age and vaccine status used in the analysis. Table 8.8 shows the influenza surveillance data and proportion of cultures positive by age and vaccine status at the time of culture. The estimates of $\mathrm{VE}^{a}_{S,IR}$, VE^{in}, and VE^{ex} are given in Table 8.9. Overall effectiveness of LAIV-T against MAARI $\mathrm{VE}^{a}_{S,IR}$ is 0.26 (95% CI: 0.11,0.39). Overall efficacy against culture-confirmed influenza using just surveillance cultures from children in the SWHP database, VE^{in}, is 0.56 (95% CI 0.24,0.84). The point estimates for VE^{in} and VE^{ex} are quite

Table 8.8. Influenza Surveillance Data (Number positive/Number cultured (proportion)), Temple-Belton, Texas, 2003-2004.

Age Group	SWHP		non-SWHP		Combined	
(years)	Unvaccinated	LAIV-T	Unvaccinated	LAIV-T	Unvaccinated	LAIV-T
5-9	8/20 (0.40)	3/15 (0.20)	19/34 (0.56)	4/9 (0.44)	27/54 (0.50)	7/24 (0.29)
10-18	35/56 (0.63)	5/13 (0.38)	30/49 (0.61)	4/11 (0.36)	65/105 (0.62)	9/24 (0.38)
Total	43/76 (0.57)	8/28 (0.29)	49/83 (0.59)	8/20 (0.40)	92/159 (0.58)	16/48 (0.33)
	TIV	PREV	TIV	PREV	TIV	PREV
5-9	2/5 (0.40)	3/9 (0.33)	0/3 (0.33)	7/21 (0.33)	2/8 (0.25)	10/30 (0.33)
10-18	3/3 (1.0)	15/29 (0.52)	5/6 (0.83)	8/15 (0.53)	8/9 (0.89)	23/44 (0.52)
Total	5/8 (0.63)	18/38 (0.47)	5/9 (0.56)	15/36 (0.42)	10/17 (0.59)	33/74 (0.44)

Table 8.9. Vaccine effectiveness of LAIV-T: $VE^a_{S,IR}$ against MAARI (95% CI), against culture-confirmed influenza using just SWHP surveillance cultures VE^{in} (95% CI), and against culture-confirmed influenza using surveillance cultures from the children in the SWHP database and children not in the SWHP database, VE^{ex} (95% CI).

Vaccine status	Age group (years)	$VE^a_{S,IR}$ (95% CI)‡ MAARI	VE^{in} (95% CI) influenza	VE^{ex} (95% CI) influenza
LAIV-T*	5-9	0.31 (0.11,0.47)	0.66 (−0.03,1.0)	0.60 (0.25,0.84)
	10-18	0.24 (0.03,0.40)	0.53 (0.12,0.86)	0.54 (0.23,0.78)
	All	0.26 (0.11,0.39)	0.56 (0.24,0.84)	0.56 (0.32,0.75)
PREV†	5-9	−0.25 (−0.61,0.05)	−0.04 (−1.9,1.0)	0.17 (−0.50,0.61)
	10-18	−0.07 (−0.28,0.10)	0.11 (−0.37,0.46)	0.09 (−0.28, 0.39)
	All	−0.13 (−0.30,0.03)	0.08 (−0.38,0.44)	0.11 (−0.19,0.37)

∗ vaccinated with LAIV-T in 2003, regardless previously vaccinated or not
† previously vaccinated in 1998-2001, but not in the 2002-2003 season or in 2003
‡ Percentile bootstrap confidence intervals based on 2000 bootstrap samples.

similar, but the confidence intervals using all of the surveillance cultures are narrower than those using just the surveillance cultures from SWHP, reflecting the higher precision conferred by the larger number of cultures.

A sensitivity analysis explored the effect of assuming that the proportion of positive cultures was constant throughout the season. The results assuming that the proportion positive varied over time were essentially identical to those in Table 8.9. Thus, assuming that the proportion of cultures that were positive was constant did not seem to influence the efficacy estimates VE^{in} and VE^{ex}.