TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

FINAL REPORT

Client:

Ministry of Health

Project Title:

Poisson Regression Modelling of the Effectiveness of the Meningococcal B Vaccine (MeNZB^{TM})

Authors:

Goals:

The main objective of this study is to develop a Poisson regression model of meningococcal disease rates in New Zealand in order to assess the impact of a new vaccine (MeNZB^{TM}). The study design of a staged roll-out of the vaccine in different age groups and geographic regions allows for an estimation of the vaccine effect that is not confounded with the progression of the epidemic over time. The Poisson regression model considers the number of subjects at risk, as well as the time at risk, so that the disease rates are calculated as the average number of cases per person per unit of time. This method of estimating rates facilitates the incorporation of the rollout vaccination schedule. The model estimates the effects of vaccination in different age and ethnic groups, while accounting for potential covariates of region-specific disease rates, deprivation, time, and seasonality.

This preliminary analysis defines cases as individuals with the epidemic strain of meningococcal disease. Other case definitions can be investigated in subsequent Poisson models. An important component of the analysis is a sensitivity study to determine the robustness of the results to various assumptions of the model and population estimates.

Executive Summary:

Several Poisson models were fit to New Zealand meningococcal case and population data. All standard Poisson models (generalised linear models) showed evidence of under-dispersion, presumably due to the correlation of observations over time and/or space. A generalised estimating equation (GEE) model approach was then used to incorporate the temporal correlations.

All models provided strong statistical evidence for a vaccine effect (p-value<0.0001). The primary analysis (based on Statistics New Zealand, SNZ, population estimates assuming medium growth and the vaccination register) estimated meningococcal disease rates to be 4.9 times higher in the unvaccinated group than in the vaccinated group (95% confidence interval = 2.7 to 9.1 times as high), with a vaccine efficacy of 80% (95% confidence interval = 63% to 89%). An estimated 75 meningococcal cases have been prevented by the vaccination programme (95% confidence interval assuming a fixed population size= 32 to 154 cases), and an estimated 2.3 deaths due to meningococcal disease have been prevented. These estimates were robust to modelling assumptions. Using SNZ low-growth population estimates, the vaccine efficacy was 81%; using SNZ high-growth estimates, the vaccine efficacy was 79%.

TE WHARE WĀNANGA O TE ŪPOKO O TE IKA A MĀUI
VICTORIA
UNIVERSITY OF WELLINGTON

Contents

1. Data Management and Model Assumptions:

1.1 Data

In this study, we assess the effect of the MeNZB^{TM} vaccine on the incidence of meningococcal cases, where the rates are defined as cases per year per 100,000 persons in the population of interest. The three main sources of data for the estimates herein are: (1) information on cases per year per demographic group provided by Ministry of Health, (2) population estimates for each year per demographic group obtained from Statistics New Zealand (SNZ) using the 2001 census as a base and projecting the population estimates to 2002-2006 assuming high, low and medium growth rates, and (3) information on the populations vaccinated per year per demographic group provided by the Ministry of Health in two forms: (a) data from the National Immunisation Register (NIR) and (b) vaccination programme start dates.

1.2 Factors affecting disease rates

The following factors are known to affect meningococcal disease rates (see for example Martin, Lopez and McDowell 2005). Rates of meningococcal disease have decreased over the period 2001-2006 in New Zealand. Furthermore, susceptibility to meningococcal disease is known to vary by age, economic status and ethnicity. To assess the effects of these factors and the efficacy of the vaccine, the following variables were included in the Poisson regression model. Age and ethnic categories, as well as seasons were chosen by grouping categories in Martin, Lopez and McDowell (2005) with similar disease rates.

(1) Year: 2001-2006.
(2) Season: defined as quarters: January – March, April – June, July – September, October – December. These quarters roughly coincide with the seasons of high and low disease rates.
(3) Region: two categorisations were considered: (A) the 21 District Health Boards (DHBs) and (B) the 4 Regional Health Authority areas (RHAs) (3 in the North Island, the fourth is the South Island).
(4) Socio-economic status: the ten NZDep01 deprivation deciles were paired to give five deprivation quintiles.
(5) Age: categorised as: less than 1 year, 1 to 4 years, 5 to 19 years, 20 or more years.
(6) Ethnicity: categorised as Maori, Pacific Peoples, European and other ethnicities.
(7) Vaccination status: not vaccinated, vaccinated

1.3 Assumptions

- This study is restricted to the time period January 2001 - June 2006.
- A meningococcal disease case is defined as an individual who has been diagnosed with the epidemic strain, verified either by PCR or by isolation.
- In the Poisson model, we only use population estimates in regions of known deprivation (NZDep01$\neq$ 0) and cases are excluded if their deprivations or ethnicities are unknown. Of 1244 cases in 2001-2006 with the epidemic strain, there were 24 cases with unknown deprivation values and 3 cases with unknown ethnicity, yielding a total of 1217 cases. All of these 27 cases were unvaccinated (in fact 26 of the 27

occurred in the years 2001-2003), so their exclusion underestimates (slightly) the meningococcal disease rates in the unvaccinated population and yields a conservative estimate of the vaccine effect. For the descriptive analysis in the next section, however, all populations and cases are counted.

- Small areas with negative or zero population estimates are excluded from analysis.
- We define vaccinated individuals as those who received three doses.
- We assume that if a case was not listed in the vaccination file, the person was not vaccinated.
- We assume that the register of vaccinations is more accurate than the SNZ population estimates, so that if the number of children vaccinated or partially vaccinated in a particular area is larger than the SNZ population estimate, then we change the population estimate to the number vaccinated plus the number partially vaccinated, and assume coverage is 100%. This assumption increased the total medium-growth scenario population estimate (for all of NZ) by 765 person-years in 2004, 4078 person-years in 2005, and 3775 person-years in 2006. These increases were less than 0.1% of the total estimate in all years.
- We assume that people vaccinated in each quarter were covered for the whole quarter.
- We assume that vaccination coverage is uniform over deprivation deciles. Available information from the NIR comprised the number of people vaccinated by quarter by DHB by ethnicity. Thus, it was necessary to estimate the number vaccinated by quarter by DHB by ethnicity by deprivation by applying each stratum's population proportion in each deprivation category to the number vaccinated in that stratum.
- Partially vaccinated individuals (those who receive one or two vaccine doses) are treated in two different ways:

  (1) assuming a potential protective effect of the partial vaccination and

  (2) assuming no protective effect.

  Under assumption (1), the population is divided into unvaccinated, partially vaccinated and fully vaccinated groups. Due to the small sample size in this group, the protective effect of partial vaccination is not estimated and the cases and population that are partially vaccinated are excluded from analysis. There were 27 partially vaccinated cases, yielding a new total of 1190 cases for this analysis. Under assumption (2), the partially vaccinated population is grouped with the unvaccinated population.

1.4 Intent to Treat Analysis

The primary analysis uses the National Immunisation Register (NIR)'s records of the number of people vaccinated in each DHB by ethnic strata in each quarter. Additionally, an intent to treat analysis was done assuming that the immunisation coverage was 100%. In this analysis, the numbers vaccinated were estimated from SNZ population estimates assuming that the vaccination programme was completed 16 weeks after the start date. The following assumptions were made in the intent to treat analysis:

- Out-of-school youth are assumed to be 18 or 19 years old.
- Subjects are fully vaccinated exactly 16 weeks after the first dose
- The population is uniformly distributed within the age categories, so there are as many 1-2 year olds as there are 2-3 year olds, for example. This assumption is necessary since the SNZ age groups do not correspond with the vaccination

programme age groups. If a vaccination programme targets 6 month to 1 year olds, we assume that 50% of this population is vaccinated and the other 50% is at risk.

Using the NIR, there were a total of 1,008,381 person-years vaccinated through June 2006. Using the intent to treat analysis, assuming 100% coverage, there were a total of 1,314,557 person-years vaccinated through June 2006.

1.5 Statistical Methodology

A Poisson rates model was suggested by Ameratunga et al. (2005) as the "most appropriate and relevant method to assess the overall effectiveness of the immunisation programme." Poisson rates models are used to model events that occur over time and space (see for example Agresti 1996 or Kianifard and Gallo 1995). We consider two types of Poisson rates models: the generalised linear model (GLIM) that assumes independent observations and the generalised estimating equation model (GEE) that allows for correlated observations over time. Both models were fit using the statistical package SAS: the GLIM model was fit using maximum likelihood, the GEE model was fit using a multivariate generalisation of quasi-likelihood.

The goodness of fit of the GLIM Poisson model was investigated by comparing the deviance to the degrees of freedom. Over- and under-dispersion were considered and two methods for adjusting for over- or under-dispersion were used: (1) adjusting standard errors with a scale statistic and (2) using a generalised estimating equation (GEE) approach. The specified correlation structure used in the GEE model was auto-regressive. In the GEE models, strata defined by region (or DHB) by ethnicity by age by deprivation by vaccination status were considered as independent populations that were observed longitudinally at 22 time points (the quarter-years of the study time period). Thus, temporal dependencies were modelled, but spatial dependencies were not incorporated into the GEE models. As discussed in Liang and Zeger (1986), however, the parameter estimates in the GEE models are robust to misspecification of the covariance structure. The GEE model thus accommodates under-dispersion, as well as misspecification in the dependency structure.

Model fitting considered all main effects and all two-way interactions. A backward step-wise procedure was used with the main effects, but a forward step-wise procedure was used with the interactions, since the model with all interactions was poorly defined and did not converge.

2. Descriptive Results:

2.1 Crude meningococcal disease rates

The crude meningococcal rates (epidemic strain) have decreased over the years 2001 – 2006 (see Table 1). A comparison of the rates between the vaccinated and unvaccinated populations shows these are similar; however, this comparison is confounded by the age of the cases, since disease rates are much higher for children and the vaccinated population consisted only of children. Figures 1-3 compare the disease rates in the vaccinated and unvaccinated populations for each age group and show a dramatic decrease in rates in the vaccinated groups (see also Table 2). A fair comparison takes account of all factors that affect disease rates, such as age, deprivation, region, etc. as in the Poisson model. The data used in the Poisson model, however, differs somewhat from that presented in Tables 1-2 and

Figures 1-3: Tables 1-2 and Figures 1-3 include all populations and cases, including those of unknown ethnicity and deprivation; the Poisson model excludes populations and cases of unknown ethnicity and deprivation and accounts for different disease rates in different regions, seasons, deprivation categories and ethnicities.

Table 1: Crude epidemic-strain meningococcal rates per year using NIR data.

	Unvaccinated			Partially vaccinated			Vaccinated		
	CASES	POPULATIO N	RATE[1]	CASES	POPULATIO N	RATE[1]	CASES	POPULATIO N	RATE[1]
2001	370	3,879,833	9.5	0	0	0	0	0	0
2002	291	3,939,240	7.4	0	0	0	0	0	0
2003	257	4,008,478	6.4	0	0	0	0	0	0
2004	178	3,993,133	4.5	7	45,577	15.4	0	29,031	0
2005	82	3,325,419	2.5	19	271,736	7.0	11	509,683	2.2
2006[2]	20	1,556,358[2]	1.3	1	43,040[2]	2.3	8	469,979[2]	1.7

[1]cases per 100,000 persons per year [2]numbers are for the half year Jan – June

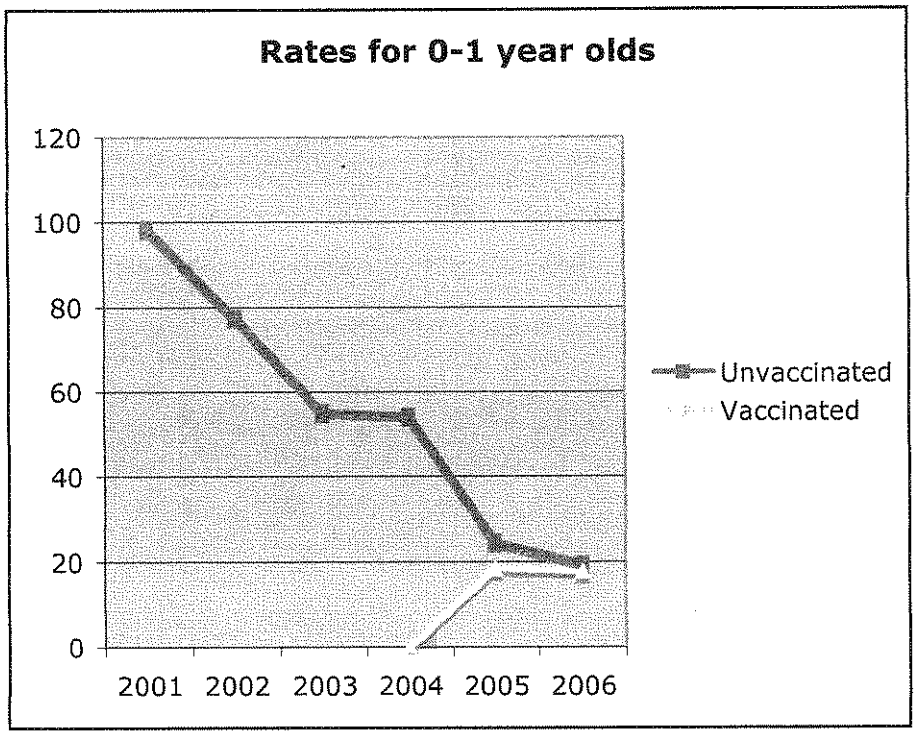

Figure 1: Crude epidemic-strain meningococcal rates per year for 0-1 year olds using NIR data.

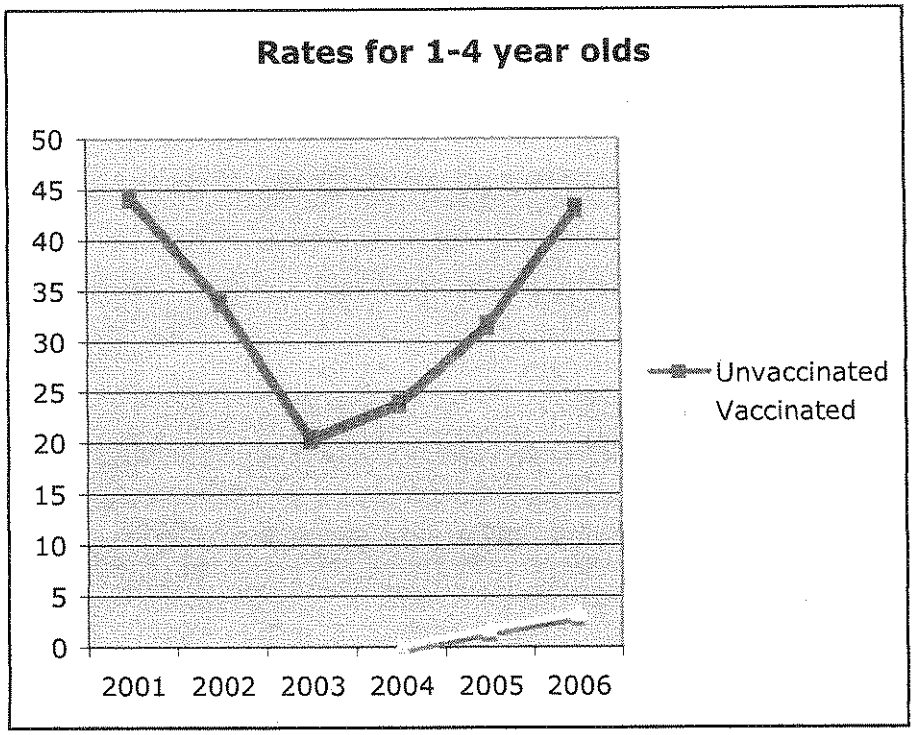

Figure 2: Crude epidemic-strain meningococcal rates per year for 1-4 year olds using NIR data.

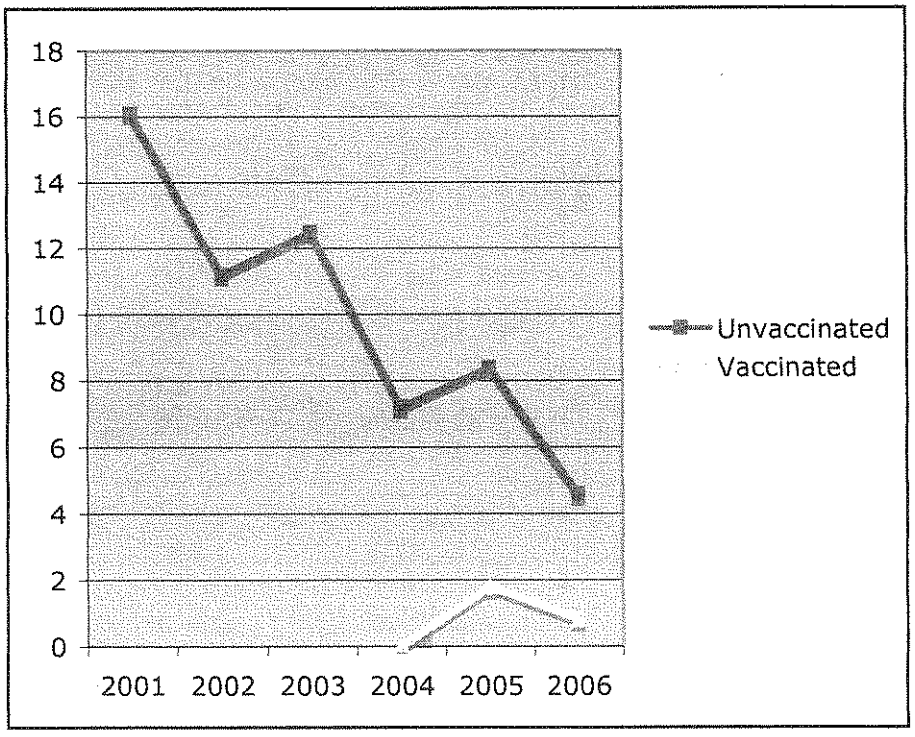

Figure 3: Crude epidemic-strain meningococcal rates per year for 5-19 year olds using NIR data.

Table 2: Crude epidemic-strain meningococcal rates per year by age using NIR data.

	Unvaccinated			Partially vaccinated			Vaccinated		
	CASES	POPULATION	RATE[1]	CASES	POPULATION	RATE[1]	CASES	POPULATION	RATE[1]
Age:	0-1 years old								
2001	55	55,805	98.6	0	0	0	0	0	0
2002	42	54,035	77.7	0	0	0	0	0	0
2003	30	54,265	55.3	0	0	0	0	0	0
2004	28	51,488	54.4	2	2,144	93.3	0	381	0
2005	6	24,315	24.7	3	18,203	16.5	2	10,911	18.3
2006[2]	1	5,100[2]	19.6	1	9,986[2]	10.0	2	11,368[2]	17.6
Age:	1-4 years old								
2001	100	225,219	44.4	0	0	0	0	0	0
2002	77	225,738	34.1	0	0	0	0	0	0
2003	46	223,651	20.6	0	0	0	0	0	0
2004	48	199,692	24.0	4	16,858	23.7	0	6,373	0
2005	18	56,589	31.8	4	47,509	8.4	2	115,685	1.7
2006[2]	3	6,923[2]	43.3	0	9,088[2]	0	3	92,335[2]	3.2
Age:	5-19 years old								
2001	141	873,743	16.1	0	0	0	0	0	0
2002	99	887,748	11.2	0	0	0	0	0	0
2003	113	900,626	12.5	0	0	0	0	0	0
2004	62	856,615	7.2	1	26,575	3.8	0	22,277	0
2005	27	320,011	8.4	12	206,025	5.8	7	383,087	1.8
2006[2]	3	64,546[2]	4.6	0	24,619[2]	0	3	366,277[2]	0.8
Age:	adults								
2001	74	2,725,066	2.7	0	0	0	0	0	0
2002	73	2,771,719	2.6	0	0	0	0	0	0
2003	68	2,829,936	2.4	0	0	0	0	0	0
2004	40	2,885,338	1.4	0	0	0	0	0	0
2005	31	2,924,503	1.1	0	0	0	0	0	0
2006[2]	13	1,479,135[2]	0.9	0	0	0	0	0	0

[1]cases per 100,000 persons per year [2]numbers are for the half year January - June

Table 3 displays the crude meningococcal rates in vaccinated, partially vaccinated and unvaccinated populations, using the intent to treat analysis assuming 100% coverage of the vaccination programme. A comparison of Tables 1 and 3 illustrates some of the differences in this methods of estimating the numbers of vaccinated people. As expected, the estimated vaccinated population is larger assuming 100% coverage making the disease rate for vaccinated individuals smaller. Additionally, the estimated unvaccinated population (the total population minus the vaccinated and partially vaccinated populations) is smaller, making the disease rate for unvaccinated individuals larger. Thus the intent to treat analysis has a larger contrast between the disease rates and the estimated vaccine effect is larger. This analysis assumes that all vaccination programmes are completed by 2006, so the partially vaccinated population has size zero in 2006.

Table 3: Crude epidemic-strain meningococcal rates per year assuming 100% coverage of the immunisation programme (intent to treat analysis).

	Unvaccinated			Partially vaccinated			Vaccinated		
	CASES	POPULATION	RATE*	CASES	POPULATION	RATE	CASES	POPULATION	RATE
2001	370	3,879,833	9.5	0	0	0	0	0	0
2002	291	3,939,240	7.4	0	0	0	0	0	0
2003	257	4,008,478	6.4	0	0	0	0	0	0
2004	178	3,987,649	4.5	7	61,027	11.5	0	19,065	0
2005	82	3,181,454	2.6	19	298,631	6.4	11	625,753	1.8
2006[2]	20	1,399,921[2]	1.3	1	0[2]	-	8	669,455[2]	1.2

*cases per 100,000 persons per year [2]numbers are for the half year January - June

2.2 Crude meningococcal mortality rates

Crude mortality rates due to the epidemic strain of meningococcal disease for the years 1999-2006 are displayed in Tables 4 and 5. There was no evidence of a change in mortality rates over the years 1999-2006 (chi-square test of independence p-value>.05), so the rates were combined to estimate an average death rate of 67/1766=3.8%. There was a significant difference in the rates across age groups (p-value=.002). When only the non-adult age groups were considered, there was still a significant difference in mortality rates (p-value=.035).

Table 4: Meningococcal death rates per year (1999-2006).

year	# of deaths	Cases	Death rate
1999	12	253	4.7%
2000	12	269	4.5%
2001	18	370	4.9%
2002	9	291	3.1%
2003	5	257	1.9%
2004	5	185	2.7%
2005	6	112	5.6%
2006	0	29	0%
Total	67	1766	3.8%

Table 5: Meningococcal death rates by age (1999-2006).

Age group	# of deaths	Cases	Death rate
0-1	14	262	5.3%
1-4	15	485	3.1%
5-19	13	629	2.1%
adult	25	390	6.4%
Total	67	1766	3.8%

3. Poisson Model Results:

Several different data sets were considered, reflecting different assumptions about the populations studied. The primary data set (on which the main results are reported) uses NIR

data and SNZ medium growth population estimates. This data set separates cases and populations into three distinct groups (vaccinated, partially vaccinated and unvaccinated), but only uses data on the vaccinated and unvaccinated cases and populations to estimate the vaccine effect. Other data sets used: (1) intent to treat analysis: vaccinated populations are estimated assuming 100% coverage of the vaccination programmes with SNZ medium growth population estimates, (2) NIR data and SNZ low-growth population estimates, (3) NIR data and SNZ high-growth population estimates, and (4) NIR data, grouping partially vaccinated populations and cases with unvaccinated populations and cases and SNZ medium growth population estimates.

We considered Poisson models that assume independence between populations in time and space (generalised linear models or GLIMs) and models that allow for correlated observations (generalised estimating equations or GEE models). In the GEE models, strata defined by RHA (or DHB) by ethnicity by age by deprivation by vaccinated status were assumed to be independent of each other, but correlated over time. As discussed in Liang and Zeger (1986), however, the parameter estimates in the GEE models are robust to misspecification of the covariance structure.

In the next section, we report the results of modelling the primary data set: SNZ medium-growth population estimates and NIR data. Then we discuss how different assumptions affected the model results.

3.1 Main Results Using Primary Data Set:

Data set 1: NIR data, excluding partially vaccinated cases and populations, and medium-growth population estimates.

All GLIM models fitted to this data set showed significant under-dispersion (deviance to degrees of freedom ratio = .42 -.45). Adjusting the covariance matrix with a scaling parameter appeared to underestimate the parameter standard errors and to be too liberal in testing the significance of model terms. Several interactions were significant in the GLIM model with an estimated scaling parameter that were not significant with the scale set to 1 or in the GEE model. Additionally, the repeated measures nature of the data suggested that a GEE model would be more appropriate.

GEE models are often used for longitudinal categorical or count data since they allow correlated observations over time: the temporal observations can be modelled as independent, exchangeable or autoregressive. We assumed the RHA (or DHB) by age by ethnicity by deprivation by vaccination strata were independent of one another, but that the temporal observations were correlated according to an AR (1) model. The estimated correlation between successive quarters in the GEE model was small (r=0.0359). Using DHBs as regional boundaries resulted in a large number of strata: 2520 strata each observed for up to 22 quarters. Since there were a total of 1190 cases, this meant that most cells had zero cases and made for computational problems (models with an interaction term involving DHB would not converge). We reduced the number of strata by using the RHA codes. This produced a total of 480 strata observed for up to 22 quarters.

Deprivation was considered as a continuous variable, and both linear and quadratic terms were investigated, but only the linear term was retained in the model. All two-way interactions were investigated.

After a backward stepwise selection procedure for the main effects in the model and a forward stepwise selection procedure for the double interaction terms, the following terms were significant in the model (see also the SAS output in Appendix A):

Table 6: Score Statistics For Type 3 GEE Analysis

Source	DF	Chi-Square	P-value
RHA	3	13.75	0.003
age	3	31.42	<0.0001
ethnicity	2	43.41	<0.0001
deprivation	1	26.83	<0.0001
year	5	38.65	<0.0001
quarter	3	59.60	<0.0001
vaccinated	1	34.70	<0.0001
deprivation*age	3	9.02	0.029
age*ethnicity	6	33.25	<0.0001
RHA*deprivation	3	10.35	0.016

3.1.1 Age and ethnicity effects:

In each ethnic group, children had significantly higher disease rates than adults (p-value<0.0001). Pacific peoples had the highest rates of meningococcal disease and this rate was especially high for young children (see Figure 4). The relative rates across age groups were very similar in the Maori and the Pacific people populations: the incidence of disease was 16-17 times higher for 1-4 year olds, 12-13 times higher for 0-1 year-olds and 4-5 times higher for 5-19 year olds than for adults. In the European and other ethnic populations, there were smaller differences in rates among the three youngest age groups: meningococcal rates were five times higher in babies (0-1 years old), seven times higher in 1-4 year olds and five times higher in 5-19 year olds than adults.

3.1.2 Deprivation effects:

Deprivation significantly affected meningococcal rates as a main effect (rates increased with increasing deprivation), and there were significant interactions between age and deprivation as well as between region (RHA) and deprivation. The effects of deprivation were significantly worse for babies: for 0-1 year olds, meningococcal rates increased 1.81 times for each quintile of deprivation (a significantly higher rate than for adults p-value=.003); for ages 1-4 years and 5 years to adults, meningococcal rates increased 1.39 and 1.36 times, respectively, for each quintile of deprivation (these two age groups were not significantly different). The effects of deprivation were the worst in the Southern RHA: in the Northern RHA, meningococcal rates increased 1.18 times for each quintile of deprivation; in the Midland RHA, meningococcal rates were stable (risk ratio=1.02) across deprivation quintiles; in the Central RHA, meningococcal rates increased 1.14 times for each quintile of deprivation; and in the Southern RHA, meningococcal rates increased 1.36 times for each quintile of deprivation.

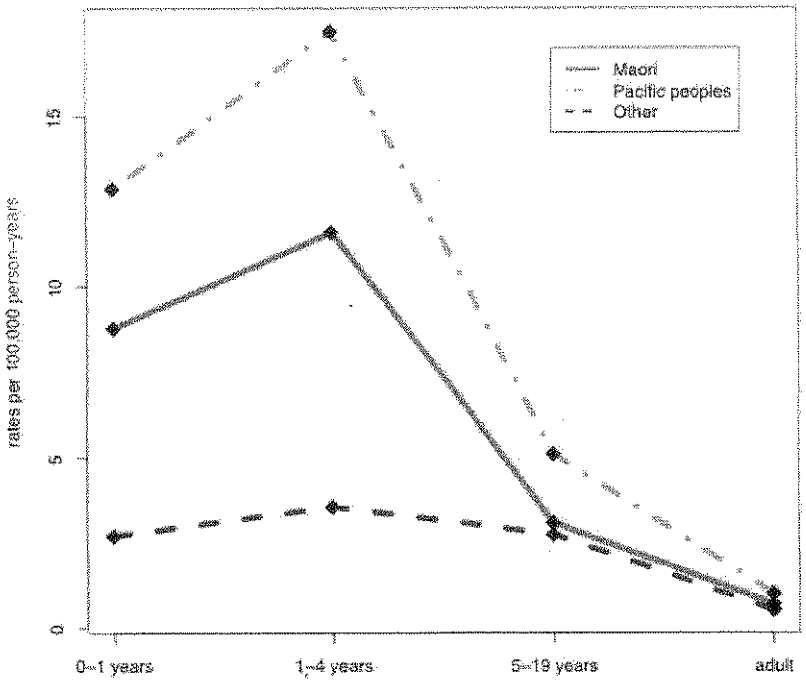

Figure 4: Estimated meningococcal rates (cases per 100,000 person-years) for each age and ethnic group in the reference category (unvaccinated persons in the lowest deprivation quintile in the third quarter of 2006).

3.1.3 Seasonal effects:

Meningococcal rates were highest in the July-September season and lowest in the January-March season (see Figure 5).

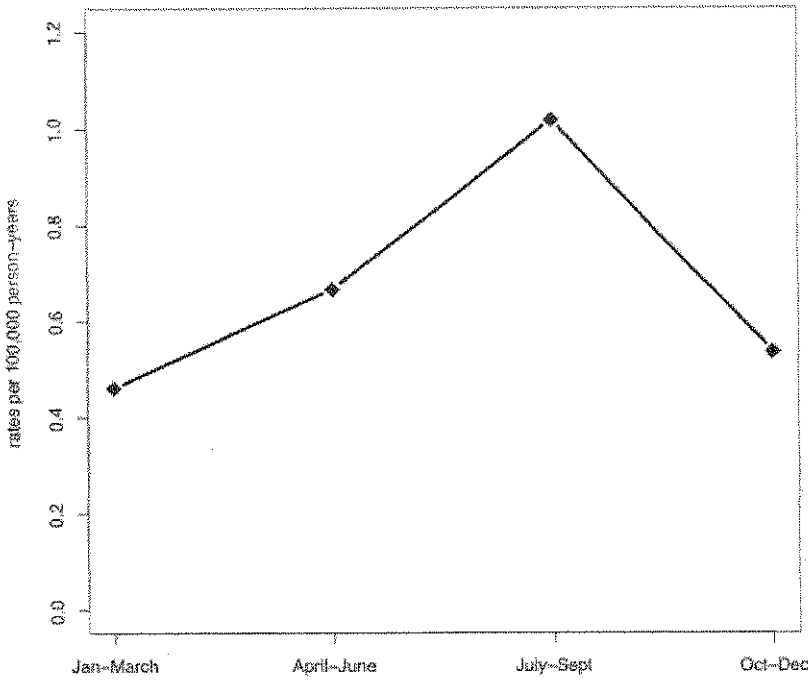

Figure 5: Estimated meningococcal rates (cases per 100,000 person-years) for each season for the reference group (unvaccinated adult Pacific people in the lowest deprivation group of European or other ethnicity in 2006).

3.1.4 Time effects:

Overall, meningococcal rates decreased over time (see Figure 6). The average decrease in disease rates from 2001-2006 was 18% per year.

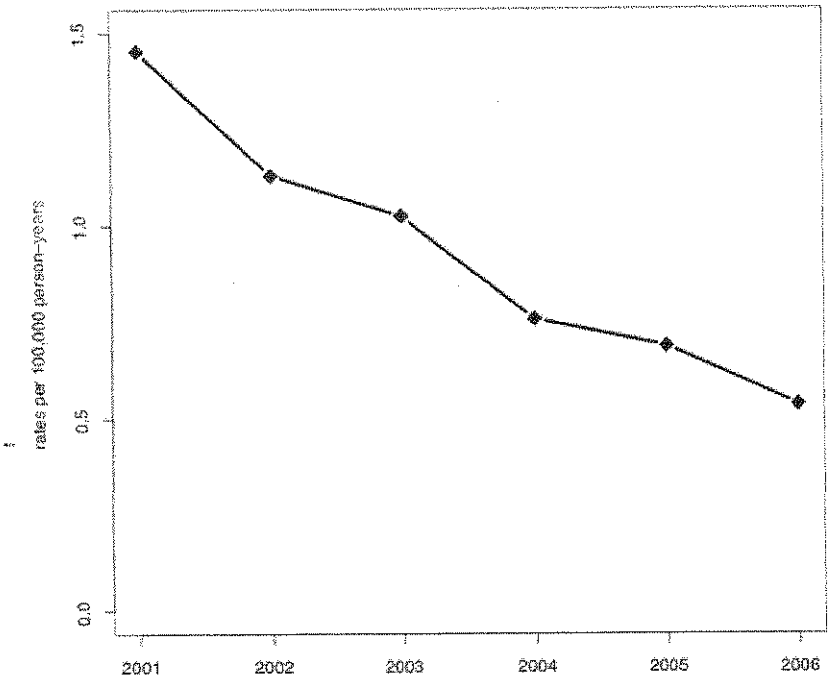

Figure 6: Estimated meningococcal rates (cases per 100,000 person-years) in the third quarter of each year for the reference group (unvaccinated adults in the lowest deprivation group of European or other ethnicity).

3.1.5 Vaccine effects:

The vaccine effect was highly statistically significant, with 4.9 times higher meningococcal disease rates in the unvaccinated group than in the vaccinated group (95% confidence interval = 2.7 to 9.1 times as high). This yields a vaccine effectiveness of (4.93-1)/4.93 = 80% (95% confidence interval = 63% to 89%).

There were a total of 1,008,693 person-years vaccinated through June 2006, according to the vaccine register, with 19 cases in this group. If this population were not vaccinated we would expect 4.93 times higher rates, or an expected 94 (rather than 19) cases. Thus, an estimated 75 meningococcal cases have been prevented by the vaccination programme (95% confidence interval assuming a fixed population size= 32 to 154 cases). Using the average mortality rate of 3.8% (for all age groups over the years 1999-2006) gives an estimated 2.8 deaths due to meningococcal disease that have been prevented. Using the age specific case numbers and mortality rates, gives an estimated 2.3 mortalities that have been prevented (see Table 7).

Table 7: Vaccinated numbers of cases and excess case and fatality numbers by age group.

Age group	# of vacc. cases	Expected # cases (if not vacc.)	Excess Cases	Death rate	Excess deaths
0-1	4	20	16	5.3%	.833
1-4	5	25	20	3.1%	.609
5-19	10	49	39	2.1%	.825
Total	19	94	75	3.8%	2.267

3.2 Robustness of the estimated vaccine effect:

We now investigate the sensitivity of the estimated vaccine effect to model assumptions, by considering alternative models and data sets.

3.2.1 GLIM Model:

Data set 1: NIR data, excluding partially vaccinated cases and populations, and medium-growth population estimates.

The GLIM fit to the primary data set had a deviance to degrees of freedom ratio of 0.45, indicating under-dispersion. The GLIM model retained the same terms as in the GEE model using a scale of 1, except that the model with an interaction term between RHA and ethnicity was favoured over the interaction term between RHA and deprivation. The model with the RHA and ethnicity interaction term did not converge for the GEE model. In both the GLIM models, the parameter estimates of common terms were similar to those of the main GEE model described above. Specifically, the vaccine effect was estimated to be 4.88 times higher rates in the unvaccinated group in the model with the RHA and ethnicity interaction term (vaccine effectiveness of 79.5%) and 4.91 times higher rates in the unvaccinated group in the model with the RHA and deprivation interaction term, with a vaccine effectiveness of 79.6%.

3.2.2 Effects Collapsing over DHBs:

Data set 1: NIR data, excluding partially vaccinated cases and populations, and medium-growth population estimates.

Models involving an interaction term with DHB did not converge. The only two significant interaction terms were age by ethnicity and age by deprivation. The vaccine effect was significant – the rate in the unvaccinated group was estimated to be 5.0 times higher than in the vaccinated group, a vaccine effectiveness of 80%.

3.2.3 Intent to treat analysis:

Data Set 2: Vaccination numbers estimated assuming 100% coverage, excluding partially vaccinated cases and populations, using medium-growth population estimates.

The intent to treat analysis yielded similar results to those presented above using NIR data. Specifically the model included all main effects and the following interactions: age by

deprivation, age by ethnicity, RHA by ethnicity and RHA by deprivation. The estimated interactions were similar to those in the primary data model reported above. Disease rates in the unvaccinated population were estimated to be 7.1 times higher than in the vaccinated population, yielding a vaccine efficacy of 86%.

3.2.4 Effects of Partial Vaccination:

Data Set 3: NIR data, grouping partially vaccinated and unvaccinated cases and populations, using medium-growth population estimates.

In the primary analysis, the partially vaccinated cases and populations were treated as a unique group, separate from the vaccinated or unvaccinated populations. Due to the small sample size in this group the protective effect of partial vaccination was not estimated, instead this group was excluded from the analysis. Here we consider how treating the partially vaccinated cases as unvaccinated affects the results. When the partially vaccinated cases and population are grouped with the unvaccinated cases and population, the same model was obtained as for the primary data set. In particular, the vaccine effect was estimated to reduce disease rates by 4.3 fold, yielding a vaccine efficacy of 76.6%. This reduction in the estimated vaccine effect indicates that there is some protective effect of partial vaccination.

3.2.5 Effects of Population Estimates:

Datasets 4 & 5: NIR data, excluding partially vaccinated cases and populations, and low and high-growth population estimates.

Using SNZ low-growth population estimates yielded the same model as the medium-growth estimates. The parameter estimates were very similar. The estimated vaccine effect was a 5.3 fold decrease in the meningococcal disease rate in the vaccinated group (95% confidence interval = 2.9-9.8), or a vaccine efficacy of 81%. This yields an estimated 82 meningococcal cases that have been prevented by the vaccination programme (95% confidence interval assuming a fixed population size= 35 to 167 cases).

Using SNZ high-growth population estimates yielded the same model as the medium-growth estimates. The parameter estimates were very similar. The estimated vaccine effect was a 4.7 fold decrease in the meningococcal disease rate in the vaccinated group (95% confidence interval = 2.6-8.7), or a vaccine efficacy of 79%. This yields an estimated 71 meningococcal cases that have been prevented by the vaccination programme (95% confidence interval assuming a fixed population size= 30 to 147 cases).

Acknowledgements

We are grateful for the contributions and previous work on the data analysis plan by David Scott, Shanthi Ameratunga and others at the University of Auckland and for the helpful comments and constructive feedback of Ministry of Health employees and peer-reviewers of this report.

References

Agresti, A. (1996). *Introduction to Categorical Data Analysis*. John Wiley and Sons, New York, NY.

Ameratunga, S., Macmillan, A., Stewart, J., Scott, D., Mulholland, K. and Crengle, S. (2005). Evaluating the post-licensure effectiveness of a group B meningococcal vaccine in New Zealand: a multi-faceted strategy. *Vaccine* 23, 2231-2234.

Kianifard, F. and Gallo, P. P. (1995). Poisson regression analysis in clinical research. *Journal of Biopharmaceutical Statistics* 5(1), 115-129.

Liang, K. Y., and Zeger, S. L. (1986). Longitudinal data analysis using generalised linear models. *Biometrika*, 73, 13-22.

Martin, D., Lopez, L. and McDowell, R. (2005). The epidemiology of meningococcal disease in New Zealand in 2004. Ministry of Health technical report.

Appendix A: SAS output for the primary data set

The GENMOD Procedure

Model Information

Data Set	WORK.MENZB_QTRQ
Distribution	Poisson
Link Function	Log
Dependent Variable	ncases
Offset Variable	lpop
Observations Used	6325

Class Level Information

Class	Levels	Values
rha	4	central midland northern southern
agecat	4	0 1 2 3
ethnicgp	3	M O P
year	6	2001 2002 2003 2004 2005 2006
qtr	4	0 1 2 3
depq	5	1 2 3 4 5
vaccinated	2	0 1

Parameter Information

Parameter	Effect	rha	agecat	ethnicgp	year	qtr	vaccinated
Prm1	Intercept						
Prm2	rha	central					
Prm3	rha	midland					
Prm4	rha	northern					
Prm5	rha	southern					

Parameter Information

Parameter	Effect	rha	agecat	ethnicgp	year	qtr	vaccinated
Prm6	agecat		0				
Prm7	agecat		1				
Prm8	agecat		2				
Prm9	agecat		3				
Prm10	ethnicgp			M			
Prm11	ethnicgp			O			
Prm12	ethnicgp			P			
Prm13	xdepq						
Prm14	year				2001		
Prm15	year				2002		
Prm16	year				2003		
Prm17	year				2004		
Prm18	year				2005		
Prm19	year				2006		
Prm20	qtr					0	
Prm21	qtr					1	
Prm22	qtr					2	
Prm23	qtr					3	
Prm24	vaccinated						0
Prm25	vaccinated						1
Prm26	agecat*ethnicgp		0	M			
Prm27	agecat*ethnicgp		0	O			
Prm28	agecat*ethnicgp		0	P			
Prm29	agecat*ethnicgp		1	M			
Prm30	agecat*ethnicgp		1	O			
Prm31	agecat*ethnicgp		1	P			
Prm32	agecat*ethnicgp		2	M			
Prm33	agecat*ethnicgp		2	O			
Prm34	agecat*ethnicgp		2	P			
Prm35	agecat*ethnicgp		3	M			
Prm36	agecat*ethnicgp		3	O			
Prm37	agecat*ethnicgp		3	P			
Prm38	xdepq*agecat		0				
Prm39	xdepq*agecat		1				
Prm40	xdepq*agecat		2				
Prm41	xdepq*agecat		3				

Prm42	xdepq*rha	central
Prm43	xdepq*rha	midland
Prm44	xdepq*rha	northern
Prm45	xdepq*rha	southern

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6294	2803.7219	0.4455
Scaled Deviance	6294	2803.7219	0.4455
Pearson Chi-Square	6294	8339.5885	1.3250
Scaled Pearson X2	6294	8339.5885	1.3250
Log Likelihood		-2074.5365	

Algorithm converged.

Analysis Of Initial Parameter Estimates

Parameter			DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept			1	-13.7889	0.4376	-14.6467	-12.9312	992.71	<.0001
rha	central		1	0.4842	0.2835	-0.0714	1.0397	2.92	0.0876
rha	midland		1	1.4681	0.2607	0.9570	1.9791	31.70	<.0001
rha	northern		1	0.4172	0.2535	-0.0797	0.9142	2.71	0.0999
rha	southern		0	0.0000	0.0000	0.0000	0.0000	.	.
agecat	0		1	2.5282	0.4853	1.5770	3.4794	27.14	<.0001
agecat	1		1	2.8360	0.3813	2.0887	3.5832	55.33	<.0001
agecat	2		1	1.6101	0.3499	0.9244	2.2958	21.18	<.0001
agecat	3		0	0.0000	0.0000	0.0000	0.0000	.	.
ethnicgp	M		1	-0.3631	0.2461	-0.8453	0.1192	2.18	0.1401
ethnicgp	O		1	-0.6402	0.2168	-1.0652	-0.2152	8.72	0.0032
ethnicgp	P		0	0.0000	0.0000	0.0000	0.0000	.	.
xdepq			1	0.3044	0.0693	0.1686	0.4401	19.31	<.0001
year	2001		1	0.9995	0.2179	0.5725	1.4266	21.05	<.0001
year	2002		1	0.7501	0.2196	0.3196	1.1805	11.66	0.0006
year	2003		1	0.6492	0.2204	0.2172	1.0812	8.68	0.0032
year	2004		1	0.3504	0.2234	-0.0875	0.7883	2.46	0.1168
year	2005		1	0.2555	0.2264	-0.1883	0.6992	1.27	0.2592
year	2006		0	0.0000	0.0000	0.0000	0.0000	.	.
qtr	0		1	-0.1507	0.0953	-0.3375	0.0360	2.50	0.1137
qtr	1		1	0.2178	0.0882	0.0450	0.3906	6.10	0.0135
qtr	2		1	0.6408	0.0818	0.4806	0.8011	61.41	<.0001
qtr	3		0	0.0000	0.0000	0.0000	0.0000	.	.
vaccinated	0		1	1.5908	0.2550	1.0911	2.0905	38.93	<.0001
vaccinated	1		0	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	0	M	1	-0.0141	0.3089	-0.6196	0.5913	0.00	0.9635
agecat*ethnicgp	0	O	1	-0.9114	0.3195	-1.5375	-0.2853	8.14	0.0043
agecat*ethnicgp	0	P	0	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	1	M	1	-0.0468	0.2860	-0.6073	0.5137	0.03	0.8701
agecat*ethnicgp	1	O	1	-0.9511	0.2789	-1.4977	-0.4044	11.63	0.0006
agecat*ethnicgp	1	P	0	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	2	M	1	-0.1376	0.2898	-0.7056	0.4305	0.23	0.6350
agecat*ethnicgp	2	O	1	0.0234	0.2612	-0.4885	0.5353	0.01	0.9286
agecat*ethnicgp	2	P	0	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	3	M	0	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	3	O	0	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	3	P	0	0.0000	0.0000	0.0000	0.0000	.	.
xdepq*agecat	0		1	0.2904	0.0921	0.1099	0.4709	9.94	0.0016
xdepq*agecat	1		1	0.0246	0.0691	-0.1109	0.1601	0.13	0.7221
xdepq*agecat	2		1	0.0043	0.0598	-0.1130	0.1216	0.01	0.9430
xdepq*agecat	3		0	0.0000	0.0000	0.0000	0.0000	.	.
xdepq*rha	central		1	-0.1731	0.0773	-0.3246	-0.0217	5.02	0.0251
xdepq*rha	midland		1	-0.2867	0.0715	-0.4268	-0.1466	16.08	<.0001
xdepq*rha	northern		1	-0.1385	0.0695	-0.2747	-0.0022	3.97	0.0464
xdepq*rha	southern		0	0.0000	0.0000	0.0000	0.0000	.	.
Scale			0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

GEE Model Information

Correlation Structure	AR(1)
Within-Subject Effect	year*qtr (22 levels)

GEE Model Information

Subject Effect	rha*age*eth*dep*vacc (418 levels)
Number of Clusters	418
Correlation Matrix Dimension	22
Maximum Cluster Size	22
Minimum Cluster Size	3

Algorithm converged.

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			-13.7994	0.4736	-14.7277	-12.8711	-29.14	<.0001
rha	central		0.4876	0.3131	-0.1261	1.1013	1.56	0.1194
rha	midland		1.4673	0.2725	0.9331	2.0014	5.38	<.0001
rha	northern		0.4204	0.2696	-0.1081	0.9488	1.56	0.1190
rha	southern		0.0000	0.0000	0.0000	0.0000	.	.
agecat	0		2.5392	0.5066	1.5463	3.5322	5.01	<.0001
agecat	1		2.8457	0.3900	2.0813	3.6102	7.30	<.0001
agecat	2		1.6143	0.3334	0.9609	2.2678	4.84	<.0001
agecat	3		0.0000	0.0000	0.0000	0.0000	.	.
ethnicgp	M		-0.3641	0.1855	-0.7277	-0.0005	-1.96	0.0497
ethnicgp	O		-0.6394	0.1752	-0.9829	-0.2960	-3.65	0.0003
ethnicgp	P		0.0000	0.0000	0.0000	0.0000	.	.
xdepq			0.3063	0.0874	0.1349	0.4776	3.50	0.0005
year	2001		0.9970	0.2346	0.5373	1.4568	4.25	<.0001
year	2002		0.7453	0.2375	0.2799	1.2108	3.14	0.0017
year	2003		0.6477	0.2213	0.2140	1.0813	2.93	0.0034
year	2004		0.3463	0.2396	-0.1234	0.8160	1.45	0.1484
year	2005		0.2490	0.2376	-0.2166	0.7147	1.05	0.2946
year	2006		0.0000	0.0000	0.0000	0.0000	.	.
qtr	0		-0.1508	0.0997	-0.3463	0.0446	-1.51	0.1304
qtr	1		0.2177	0.0789	0.0630	0.3724	2.76	0.0058
qtr	2		0.6414	0.0797	0.4851	0.7977	8.04	<.0001
qtr	3		0.0000	0.0000	0.0000	0.0000	.	.
vaccinated	0		1.5961	0.3125	0.9836	2.2087	5.11	<.0001
vaccinated	1		0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	0	M	-0.0185	0.2579	-0.5240	0.4869	-0.07	0.9427
agecat*ethnicgp	0	O	-0.9175	0.3247	-1.5540	-0.2811	-2.83	0.0047
agecat*ethnicgp	0	P	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	1	M	-0.0473	0.2334	-0.5048	0.4101	-0.20	0.8392
agecat*ethnicgp	1	O	-0.9513	0.2650	-1.4707	-0.4319	-3.59	0.0003
agecat*ethnicgp	1	P	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	2	M	-0.1351	0.2470	-0.6192	0.3489	-0.55	0.5843
agecat*ethnicgp	2	O	0.0238	0.2229	-0.4130	0.4606	0.11	0.9149
agecat*ethnicgp	2	P	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	3	M	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	3	O	0.0000	0.0000	0.0000	0.0000	.	.
agecat*ethnicgp	3	P	0.0000	0.0000	0.0000	0.0000	.	.
xdepq*agecat	0		0.2885	0.0979	0.0967	0.4803	2.95	0.0032
xdepq*agecat	1		0.0223	0.0733	-0.1214	0.1661	0.30	0.7607
xdepq*agecat	2		0.0028	0.0631	-0.1209	0.1266	0.05	0.9641

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

Parameter		Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
xdepq*agecat	3	0.0000	0.0000	0.0000	0.0000	.	.
xdepq*rha	central	-0.1737	0.0948	-0.3596	0.0121	-1.83	0.0669
xdepq*rha	midland	-0.2862	0.0834	-0.4497	-0.1228	-3.43	0.0006
xdepq*rha	northern	-0.1391	0.0832	-0.3022	0.0240	-1.67	0.0946
xdepq*rha	southern	0.0000	0.0000	0.0000	0.0000	.	.

Score Statistics For Type 3 GEE Analysis

Source	DF	Chi-Square	Pr > ChiSq
rha	3	13.75	0.0033
agecat	3	31.42	<.0001
ethnicgp	2	43.41	<.0001
xdepq	1	26.83	<.0001
year	5	38.65	<.0001

qtr	3	59.60	<.0001
vaccinated	1	34.70	<.0001
agecat*ethnicgp	6	33.25	<.0001
xdepq*agecat	3	9.02	0.0290
xdepq*rha	3	10.35	0.0158

Appendix B: DHB Codes

DHB code	DHB name
011	Northland District Health Board
021	Waitemata District Health Board
022	Auckland District Health Board
023	Counties Manukau District Health Board
031	Waikato District Health Board
042	Lakes District Health Board
047	Bay of Plenty District Health Board
051	Tairawhiti District Health Board
061	Taranaki District Health Board
071	Hawkes Bay District Health Board
081	MidCentral District Health Board
082	Whanganui District Health Board
091	Capital and Coast District Health Board
092	Hutt Valley District Health Board
093	Wairarapa District Health Board
101	Nelson Marlborough District Health Board
111	West Coast District Health Board
120	Canterbury District Health Board
123	South Canterbury District Health Board
131	Otago District Health Board
141	Southland District Health Board

Appendix C: SAS Code

File 1: fcmd01_readdata.sas

This program reads in fifteen data files containing: (1) population estimates, (2) case information, (3) vaccination history of selected cases, (4) start dates for the vaccination programme, (5) population estimates for Auckland EC, (6-9) vaccination and partial vaccination coverage rates for 2004 and 2005, and (10-15) high, low and medium population estimates. A minor amount of data processing is done, including the following.

- DHBs are coded into their 3 number codes.
- Ethnicity is recoded as a one-letter code.
- Negative population estimates are recoded to zero.
- Areas with nzdep01 of 0 are discarded.
- Only cases meeting the Epidemic strain case definition, with known ethnicity, known nzdep01, and with onset between 2001-2005 are retained.

File 2a: fcmd02_prepare.sas (Number Vaccinated Estimation Method I)

This program uses the number of people vaccinated and partially vaccinated provided in the vaccination register. People vaccinated in a particular quarter are assumed to be protected for the entire quarter. Partially vaccinated people are discarded from the model. The data processing consists of the following steps:

Step 1: *Recode age groups in the SNZ population estimates file* (variable `agegrp`) *to correspond with the four age categories used in the Poisson model* (variable `agecat`).

Step 2: *Estimate the numbers vaccinated and partially vaccinated in Counties Manakau and Auckland EC (these are combined in the vaccination register), by assuming the vaccinated proportions in each region are equal to their relative population sizes.*

Step 3: *Estimate the numbers vaccinated in each deprivation category (these are not provided in the vaccination register), by assuming the same proportion is vaccinated in each deprivation category.*

Step 4: *Calculate the numbers not vaccinated in each DHB by ethnicity by age category by deprivation by quarter group as the total population minus the number vaccinated minus the number partially vaccinated.*

Step 5: *Ensure that cases meet the case definition and recode the age groups, then merge with the population file. Regions with zero or negative population estimates are discarded.*

File 2b: fcmd02q_fracvac.sas (Number Vaccinated Estimation Method II)

This program uses the vaccination programme start dates to calculate the fraction of each dhb by agegroup by quarter that is vaccinated, assuming 100% coverage of the vaccination programs. The data processing consists of the following steps:

Step 1: *Split age groups into yearly quarters, so that each quarter is covered by one and only one vaccination programme.*
This program splits the SNZ population age groups into quarters, so that each new age group (variable `age`) is covered by only one vaccination programme. The file temp contains the original age groups given in the population estimates file (variable `agegrp`), the vaccination programme start dates (variables `dprimary, dschool, d6wk5m`), and a new vaccination age group variable which links the vaccination programme age groups to the population estimate age groups (variable `vagegrp`). The number of years in each age group is stored in the variable `yrs_agegrp`. All these variables are stored in the data file vstart_qtr.

Step 2: *Calculate the fraction of each quarter covered by a vaccination programme (fully vaccinated) as well as the fraction of the quarter that is partially vaccinated for each dhb by age by quarter category.*
This program first calculates the estimated finish times (variables `fdate_primary, fdate_school, fdate_6wk5m`) for each vaccine programme by adding 16 weeks to

the start date. Then the proportion of the quarter that each dhb by age category was fully covered by the vaccine is calculated (variable `frac_qtr`), as well as the proportion of the quarter that was partially vaccinated (variable `pfrac_qtr`). Finally, each age and quarter combination are assigned to a cohort, so that previously vaccinated age groups can be rolled into successive quarters. These variables are saved in the data file xvstart_qtr.

Step 3: *Roll previously vaccinated people into the next quarter.*
We assume here that once a vaccination programme begins that it continues indefinitely, so that all children in that age group will be vaccinated after the programme's start date. If a programme has not yet started, it is possible to have some children who were vaccinated by another programme that started at an earlier date. For example, if the 6wk- 5month programme starts the first quarter of 2004, then these children will be in the 1-4 year old age group in 2005 and they will be vaccinated. The total fraction vaccinated in a quarter cohort is thus those who were previously vaccinated plus those who are newly vaccinated, each of which is weighted by the proportion of the quarter that they were vaccinated:.

```
pqtrs_vacc =    frac_qtr*prop_cohort_new_vacc +
                (1-frac_qtr)*prop_cohort_already_vacc;
```

Step 4: *Average over age groups to obtain the vaccinated and partially fraction in each age category by quarter.*

File 3: fcmd03q_prepare.sas (Number Vaccinated Estimation Method II)

This file prepares the data in a manner analogous to File 2a.

File 4-5: fcmd02nq_high_prepare.sas **and** ***fcmd02nq_low_prepare.sas*** are the same as File 2a (***fcmd02nq_prepare.sas)***, but they use the high and low SNZ population estimates.

File 6: fcmd02nq_pvwithuv_prepare.sas is similar to ***fcmd02nq_prepare.sas***, except that it combines the partially vaccinated population with the unvaccinated population.

File 7: recodedhb.sas recodes the DHBs into RHAs and NZ deprivation deciles into quintiles and then runs the GEE and GLIM models.