Epidemiol. Infect. (2003), **130**, 201–205. © 2003 Cambridge University Press
DOI: 10.1017/S0950268802008130 Printed in the United Kingdom

Analysis of longitudinal bacterial carriage studies accounting for sensitivity of swabbing: an application to *Neisseria meningitidis*

C. L. TROTTER* AND N. J. GAY

Immunisation Division, PHLS Communicable Disease Surveillance Centre, 61 Colindale Avenue, London, NW9 5EQ, UK

(*Accepted 31 October 2002*)

SUMMARY

Longitudinal carriage studies of colonizing bacteria such as *Neisseria meningitidis* can provide important insights into the transmission dynamics of these organisms. Carriage is detected by culturing from a nasopharyngeal swab, but the sensitivity of this technique is low and varies between studies. This paper applies a statistical method for estimating the sensitivity of swabbing, infection rate, recovery rate and initial prevalence of carriage to three longitudinal carriage studies of *N. meningitidis*. These parameters and 95% confidence intervals were estimated using maximum likelihood techniques. The sensitivity of swabbing was estimated to be 60–83% and this should be taken into account when interpreting carriage studies. The estimates of force of infection and recovery rates seem to be consistent with estimates from more traditional methods. Differences in the parameter estimates between datasets may be due to differences in study design. This method could be used to assist in the design of future carriage studies.

INTRODUCTION

Colonizing bacteria such as *Neisseria meningitidis*, *Haemophilus influenzae* and *Streptococcus pneumoniae* are frequently present in the nasopharyngeal flora of healthy persons [1–3]. Invasive disease is a rare outcome of infection and asymptomatic carriers are responsible for most transmission of these organisms. Studies of carriage, detected by culturing samples from a nasopharyngeal swab, can therefore provide important insights into the transmission dynamics of these organisms. Longitudinal carriage studies, in which individuals are swabbed repeatedly over several weeks or months, can be particularly informative. The specificity of swabbing is almost 100%, but the sensitivity of this technique may be as low as 50% [4]. Longitudinal studies usually account for low sensitivity by applying somewhat arbitrary rules to define conversion between carrier and non-carrier status for estimating recovery and infection rates. To eliminate the need for such assumptions a statistical method for estimating the force of infection (λ), recovery rate (ν), initial prevalence (π) and sensitivity of swabbing (σ) from longitudinal data was developed, following the work of Nagelkerke et al. [5].

This method is applied to data from two longitudinal studies of *N. meningitidis* carriage in military recruits [4, 6] and currently unpublished data from a community carriage study in Stonehouse, Gloucester.

METHODS

Model

The method is designed to analyse longitudinal data, in which individuals are tested at common times to generate a series of test results for each individual. The test result sequence of an individual is a function of the sensitivity of the test (σ) and the true status sequence (whether an individual is truly a carrier or

* Author for correspondence.

not). The true status sequence is a function of parameters λ (the force of infection), ν (the recovery rate) and π (the initial prevalence). Parameters σ, λ, ν and π are estimated for the whole dataset using maximum likelihood methods.

The model of the true status sequence

Each individual was assumed to exist in one of two true states, carrier (C, proportion c) or susceptible (S, proportion $s=1-c$). Carriers were assumed to recover at a rate ν and susceptibles to become infected at a rate λ. According to the standard SIS model,

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -\lambda s + \nu c$$

$$\frac{\mathrm{d}c}{\mathrm{d}t} = \lambda s - \nu c,$$

the proportion of susceptibles and carriers at time t depends on s_0 and c_0, the respective proportions at time 0,

$$s(t;s_0) = \frac{\nu}{\lambda+\nu}(1-\mathrm{e}^{-(\lambda+\nu)t}) + s_0 \cdot \mathrm{e}^{-(\lambda+\nu)t}, \tag{1}$$

$$c(t;c_0) = \frac{\lambda}{\lambda+\nu}(1-\mathrm{e}^{-(\lambda+\nu)t}) + c_0 \cdot \mathrm{e}^{-(\lambda+\nu)t}. \tag{2}$$

The transition probabilities of moving between states in a time interval Δt [5] can be derived from equations (1) and (2) as follows:

$$\begin{aligned}\mathrm{pr}(S \to C{:}\Delta t) &= c(\Delta t; c_0 = 0)\\ &= \frac{\lambda}{\lambda+\nu}(1-\mathrm{e}^{-(\lambda+\nu)\Delta t})\end{aligned} \tag{3}$$

$$\begin{aligned}\mathrm{pr}(S \to S{:}\Delta t) &= s(\Delta t; s_0 = 1)\\ &= \frac{\nu}{\lambda+\nu}(1-\mathrm{e}^{-(\lambda+\nu)\Delta t}) + \mathrm{e}^{-(\lambda+\nu)\Delta t}\end{aligned} \tag{4}$$

$$\begin{aligned}\mathrm{pr}(C \to S{:}\Delta t) &= s(\Delta t; s_0 = 0)\\ &= \frac{\nu}{\nu+\lambda}(1-\mathrm{e}^{-(\lambda+\nu)\Delta t})\end{aligned} \tag{5}$$

$$\begin{aligned}\mathrm{pr}(C \to C{:}\Delta t) &= c(\Delta t; c_0 = 1)\\ &= \frac{\lambda}{\lambda+\nu}(1-\mathrm{e}^{-(\lambda+\nu)\Delta t}) + \mathrm{e}^{-(\lambda+\nu)\Delta t}.\end{aligned} \tag{6}$$

In a data set with N test points, there are 2^N permutations of status sequence. Let Q_{in} denote the status at the nth test point of the ith status sequence ($Q_{in} = S$ or C) and Y_{in} denote the probability of the nth transition in the ith status sequence. At the first test point Y_{i1} depends upon π, the initial prevalence. At all subsequent test points the true status depends on the transition probabilities described in equations (3)–(6) above. So,

$$Y_{i1} = \mathrm{pr}(Q_{i1}) \begin{pmatrix} = \pi & \text{if } Q_{i1} = \mathrm{C} \\ = 1-\pi & \text{if } Q_{i1} = \mathrm{S} \end{pmatrix}$$

and

$$Y_{in} = \mathrm{pr}(Q_{i,n-1} \to Q_{i,n}), \quad n = 2, \ldots, N.$$

The probability of the ith status sequence is denoted $W_i(\pi, \lambda, \nu)$. W_i is equal to the product of the probability of the initial status and the transition probabilities at each test point.

$$W_i\ (\pi, \lambda, \nu) = \prod_{n=1}^{N} Y_{in}.$$

Test results

At each test point an individual was classified as a positive, negative or not tested for the presence of the colonizing organism. Data on the strain characteristics (serogroup and serotype) of the infecting organism were not taken into account.

The model of the observations, given the underlying true status sequence

At any occasion on which a test is performed, the probability of observing a particular test result depends on the true status of the individual and on the sensitivity, σ, of the test.

Hence, the probability V_{ij} of observing the test results of the jth individual if they truly have the ith status sequence is,

$$V_{ij} = \sigma^{d_{ij}} \cdot (1-\sigma)^{e_{ij}} \cdot 0^{f_{ij}}, \tag{8}$$

where:

d_{ij} = number of times test is positive when true status is carrier

e_{ij} = number of times test is negative when true status is carrier

f_{ij} = number of times test is positive when true status is susceptible

We assumed that the specificity of the test was 100%, i.e. it was not possible to test positive when the true status was susceptible, so if $f_{ij} > 0$, then $V_{ij} = 0$. The calculation of d_{ij}, e_{ij}, and f_{ij} for each individual (j) and status sequence (i) is independent of the parameters σ, λ, ν, π.

https://doi.org/10.1017/S0950268802008130 Published online by Cambridge University Press

Table 1. *Summary of data sources*

	Stonehouse	Riordan	Pether
Setting	Community	Military	Military
Year of study	1986/7	1994/5	1986/7
Length of study period	14 months	29 weeks	14 months
No. individuals in study	436	311	283
No. individuals included in this analysis	85	255	226
Median number of swabs per individual	3	7	4
Reference	K. Cartwright and J. Stuart (unpublished)	[6]	[4]

Parameter estimation

The overall probability $Z_j(\pi, \lambda, \nu, \sigma)$ of observing the results for individual j is thus

$$Z_j(\pi,\lambda,\nu,\sigma) = \sum_{i=1}^{2^N} W_i\ (\pi,\lambda,\nu) \cdot V_{ij}(\sigma). \tag{9}$$

We assumed that each individual is independent, so that the overall log likelihood L is

$$L = \sum_j \log Z_j = \sum_j \log\left(\sum_{i=1}^{2^N} W_i \cdot V_{ij}\right). \tag{10}$$

The parameters σ, π, λ and ν were estimated by maximizing the log likelihood L. The Confidence Profile Method [7] was used to approximate 95% confidence intervals for the four parameters, by minimizing and maximizing (for lower and upper CI respectively) the values of each parameter for which L was equal to the maximum likelihood $-1{\cdot}922$. Analyses were conducted both in Microsoft Excel and using a standard optimization procedure (Powell's method [8]) in Fortran. This approach to optimizing the likelihood differs from that of Nagelkerke [5]. We were able to use this more straightforward method because individuals in our datasets were swabbed at common test points and there were fewer observations per individual, with a maximum number 14 (compared to 44 in Nagelkerke's data).

Data

This method was applied to data from three longitudinal carriage studies of *N. meningitidis*, where the data were published or made available by the authors (Table 1). Pether et al. [4] studied personnel on a Royal Naval Air Station in 1986/7. The data comprised the results of swabs from 283 recruits swabbed repeatedly over a year on up to 14 occasions. Any individual swabbed on only one occasion during the study period was excluded, because it was not possible to observe any loss or acquisition of carriage in these individuals. Data from 226 individuals was analysed. Riordan et al. [6] studied the acquisition and carriage of meningococci in marine commando recruits in 1994/5. Individuals were swabbed on up to eight occasions over the study period, which lasted 29 weeks. Because the study period and the interval between swabs in this study was shorter than Pether's study, individuals with two or fewer swabs were excluded leaving the results from 255 individuals for analysis. Recruits in this study were swabbed at eight different time points; however, a small number of recruits who had been 'back-trooped' had 9 or 10 swabs taken. When this occurred, the result of the first swab taken in that week was retained and the second test result of the week was deleted.

The Stonehouse data was collected as part of a larger community wide survey, following an outbreak of group B meningococcal disease in 1986/7 [1]. These data were kindly provided by Professor Keith Cartwright and Dr James Stuart. Individuals were invited to participate in the longitudinal study if they had been previously identified as a carrier, or if they were in the same household as a carrier. Results were analysed from the 85 individuals who were followed for up to 14 months, swabbed on more than two occasions and were not treated with rifampicin. Because individuals were selected on the basis of their carrier status or relationship with a carrier the first testing occasion could not contribute to the calculation of sensitivity. To account for this, tests at time t_0 were not included in the calculation of d_{ij}, e_{ij}, f_{ij}.

RESULTS

The results of these analyses are shown in Table 2. The sensitivity of swabbing is highest in Riordan's study at around 83%, but is lower in both Pether's study (60%) and the Stonehouse study (65%). The highest force of infection is seen in Riordan's study of marine

https://doi.org/10.1017/S0950268802008130 Published online by Cambridge University Press

Table 2. *Estimates of sensitivity, infection rate* (*per week*), *recovery rate* (*per week*), *the duration of carriage* (*months*) *and initial prevalence in three data sets*

Data set	Sensitivity of swabbing σ (95% CI)	Force of infection λ (95% CI)	Recovery rate ν (95% CI)	Mean duration of carriage $1/\nu$ (in months) (95% CI)	Initial prevalence π (95% CI)
Pether	59·5% (53·3–65·7)	0·003 (0·001–0·006)	0·008 (0·003–0·015)	29 (15–77)	35% (28–43)
Riordan	83·4% (78·2–87·0)	0·048 (0·04–0·06)	0·040 (0·028–0·070)	6 (3–8)	39% (33–46)
Stonehouse	65·2% (57·2–73·2)	0·0001 (0–0·011)	0·011 (0·006–0·018)	21 (13–39)	99% (91–100)

recruits and lowest in the Stonehouse study. The mean duration of carriage varies from 6 months in Riordan's study up to 21 and 29 months in Stonehouse and Pether's respectively. The median duration of carriage (median = ln(2)* mean) is estimated to be 14·5 months for Stonehouse, 4 months for Riordan's study and 20 months for Pether's study.

These results can be compared with those derived through more traditional methods. For example, the force of infection as estimated from Riordan's study indicates that approx. 17% of the remaining susceptibles would acquire a meningococcus over any month. This compares favourably with the data published by Riordan et al., which reports an acquisition rate of 15–25% per month. The Stonehouse data were analysed using the 'evaluation of truncated observations' method [9], and a mean duration of 7·1 months was estimated. This is much lower than estimated using the model, and probably reflects the high proportion of individuals with truncated observations. The original paper by Pether et al. does not estimate the force of infection or recovery rate but they did estimate the sensitivity of swabbing and found that it may have been as low as 50%. This is lower than estimated using the present methods.

DISCUSSION

The sensitivity of swabbing is estimated to be well below 100% in each of the studies, ranging between 60 and 83%. The differences between the studies may be attributable to differences in microbiological techniques and the skill of the swabbers [10]. This low sensitivity must be taken into account when interpreting the results of any carriage study. Since the sensitivity is not uniform and varies between studies it would be desirable for future carriage studies to incorporate a parallel sensitivity study (for example comparing the isolation rate from different swabbers), especially for cross-sectional carriage studies to which this statistical method cannot be applied.

The estimates of the infection rate are also quite variable, but this could be explained by differences in the study populations. The higher rate of infection in Riordan's study compared to Pether's study could be explained by more intense contact in new military recruits compared to established troops. In the case of the Stonehouse data, individuals were selected for the study on the basis of their carrier status, so the infection rate is low. A lower infection rate would be expected regardless in a community setting compared to a military setting. However, the rate of reinfection within those individuals clearing infection may be lower if infection generates some immunity.

An initial prevalence of 35 and 39% was estimated for the data of Pether et al. and Riordan et al. These carriage rates are high, but not unusually so for military establishments. The initial prevalence in the Stonehouse data reflects the selection of subjects on the basis of their carrier status.

One would however, expect the estimates of the recovery rate to be more similar, since this is biologically determined. In our model the recovery rate is assumed to be constant with an exponential distribution of carriage duration. Other studies have estimated that the median duration of meningococcal carriage is between 7·7 and 10·2 months, using techniques such as survival curves [9, 11]. Unfortunately these data were not available for analysis using this method, but the median durations of carriage estimated here are not too far removed from this range. The median duration of carriage in Pether et al.'s study and the Stonehouse study are 20 months and 14·5 months respectively. The mean duration of carriage estimated from Riordan et al. is shortest, at 6 months (median 4 months), and the highest mean duration of carriage within the 95% confidence interval is 8 months. While the recovery rate may vary according to age and the number of previous infections, this is unlikely to explain the differences between Pether's and Riordan's data.

Differences in study design, particularly differences in the sampling interval and number of tests per individual, may contribute to the differences in parameter estimates. Riordan's study, which had the highest average number of swabs per individual and the shortest sampling interval, yielded the highest sensitivity, force of infection and recovery rate estimates. Using the method described here to analyse simulated datasets could assist in the design of future carriage studies. For example, different sampling intervals could be simulated and the advantages of increasing the number of tests per individual *vs*. increasing the number of individuals in the study could be investigated.

There is some correlation between parameter estimates. Sequences of positive results separated by negative results can either be explained by lack of sensitivity or by true conversions. Thus high sensitivity estimates will correlate with high force of infection and recovery rates.

The model classifies individuals as a carrier (or not) and there is no differentiation between carriage of different strains, and the total force of infection and average recovery rate for all meningococci are measured. Several individuals in each dataset are shown to carry more than one strain during the study period. This could be due to an individual clearing one strain and acquiring another, or it could be that those individuals were carrying both strains simultaneously, and multiple carriage is not detected because only one colony is sampled at each test point. Incorporating strain characteristics would make the model much more complex and difficult to interpret, so the simpler binary carrier-susceptible model is preferred.

The model further assumes that the rate of acquisition is constant throughout the study. Ideally the model should be dynamic, and take into account changes in the number of infected individuals in the population. However, changes in the prevalence of infection during the course of the studies are modest so this would complicate the model for possibly minimal gains in precision.

Individuals are assumed to be susceptible to reinfection as soon as an episode of carriage has ended, and there is no period of immunity following carriage. If individuals are not susceptible for a period of weeks or months after clearing carriage then the force of infection will be under-estimated here. Immunological studies have shown that bacterial carriage stimulates an immune response against the infecting organism [12]. However, while this is likely to confer protection against invasive disease, repeated episodes of carriage are common and there is little information on the duration of protection following natural infection. It seems reasonable to assume here that immunity to carriage is short lived, so it is not incorporated into the model.

Despite these limitations of the study, we feel it provides useful insights into meningococcal dynamics and may be valuable in assisting with the design of any future longitudinal studies.

ACKNOWLEDGEMENTS

We thank Keith Cartwright, James Stuart and Terry Riordan for providing data to make this study possible. We also thank Keith Cartwright, James Stuart, Kari Auranen, Andrew Grant, Nick Andrews and an anonymous referee for helpful comments and suggestions.

REFERENCES

1. Cartwright KA, Stuart JM, Jones DM, Noah ND. The Stonehouse survey: nasopharyngeal carriage of meningococci and *Neisseria lactamica*. Epidemiol Infect 1987; **99**: 591–601.
2. Howard AJ, Dunkin KT, Millar GW. Nasopharyngeal carriage and antibiotic resistance of *Haemophilus influenzae* in healthy children. Epidemiol Infect 1998; **100**: 193–203.
3. Ghaffar F, Friedland IR, McCracken GH. Dynamics of nasopharyngeal colonization by *Streptococcus pneumoniae*. Pediatr Infect Dis J 1999; **18**: 638–46.
4. Pether JV, Lightfoot NF, Scott RJ, Morgan J, Steele-Perkins AP, Sheard SC. Carriage of *Neisseria meningitidis*: investigations in a military establishment. Epidemiol Infect 1988; **101**: 21–42.
5. Nagelkerke NJD, Chunge RN, Kinoti SN. Estimation of parasitic infection dynamics when detectability is imperfect. Stat Med 1990; **9**: 1211–9.
6. Riordan T, Cartwright K, Andrews N, et al. Acquisition and carriage of meningococci in marine commando recruits. Epidemiol Infect 1998; **121**: 495–505.
7. Eddy DM. The confidence profile method: a Bayesian method for assessing health technologies. Oper Res 1989; **37**: 210–28.
8. Teukolsky SA, Vettering WT, Flannery BP. Numerical recipes. Cambridge: Cambridge University Press, 1986.
9. De Wals P, Bouckaert A. Methods for estimating the duration of bacterial carriage. Int J Epidemiol 1985; **14**: 628–34.
10. Broome CV. The carrier state: *Neisseria meningitidis*. J Antimicrob Chemother 1986; **18** (Suppl A): 25–34.
11. Greenfield S, Sheehe PR, Feldman HA. Meningococcal carriage in a population of 'normal' families. J Infect Dis 1971; **123**: 67–73.
12. Goldschneider I, Gotschlich EC, Artenstein MS. Human immunity to the meningococcus. II. Development of natural immunity. J Exp Med 1969; **129**: 1327–48.