
Transmission and distribution of virus serotypes: African horse sickness in zebra

C. C. LORD¹*, M. E. J. WOOLHOUSE¹ AND B. J. H. BARNARD²

¹ Department of Zoology, South Parks Road, University of Oxford, Oxford OX1 3PS

² Onderstepoort Veterinary Institute, Private Bag X5, Onderstepoort 0110, Republic of South Africa

(Accepted 11 June 1996)

SUMMARY

The prevalence of African horse sickness (AHS) serotypes in zebra foals from the Kruger National Park, South Africa was examined for possible associations between serotypes and to estimate the basic reproduction number, R_0 . The distributions of serotypes between zebra were not independent in the 6- and 7–8-month-old age classes ($P < 0.005$). This does not necessarily imply biological interactions between serotypes, as heterogeneity in host-vector transmission rates can generate non-independent distributions of serotypes. Both age and month of capture were significant factors in the number of serotypes infecting each zebra ($P < 0.0001$). Pairwise, positive associations between non-cross-reacting serotypes were found in the 7–8-month-old class only. For AHS overall, estimates of R_0 ranged from 31–68. Assuming serotypes are transmitted independently, estimates of R_0 for individual serotypes ranged from 10 for serotype 1 to 23 for serotype 6. The wide range of estimates emphasizes the need for a better understanding of serotype transmission and interactions in AHS.

INTRODUCTION

Multiple strains or serotypes have been found in several parasite species, including protozoan diseases such as malaria [1] and trypanosomiasis [2], and viruses (e.g. African horse sickness [3], bluetongue [4], foot and mouth [5], dengue [6]). For convenience, we will refer only to serotypes although the methods used in this paper are equally applicable to parasites other than viruses. Only recently have there been attempts to understand the epidemiological implication of serotypes, but it is clear that there are consequences for the basic reproduction number, R_0 [7]. This number is a measure of the strength of transmission of a parasite or pathogen, and is defined as the average number of secondary cases generated by one primary case in a susceptible population. R_0 can be used in the design of control campaigns; for example, to indicate the coverage required in a vaccination program [8]. It

has been suggested by Gupta and colleagues [7] that conventional methods of calculating R_0 may overestimate the true value if the disease is a composite of several independently transmitted serotypes, and they have suggested using a weighted average of R_0 in this case.

The magnitude of the impact on R_0 depends on the extent to which the serotypes are transmitted independently. Independent transmission here is defined as transmission without biological interactions between the serotypes. Transmission would not be independent if infection with one serotype affected the potential for infections with another serotype; for example, by inducing a cross-reactive immune response. It is important to distinguish independent transmission of serotypes from the independent distribution of serotypes among hosts. The latter may be generated by heterogeneities in transmission rates between hosts which need not imply interactions between serotypes. The detection of non-independent distributions of serotypes is analogous to the problem

* Current address: Florida Medical Entomology Laboratory, 200 9th Street SE, Vero Beach, FL 32962, USA.

of determining independence between species distributions in community ecology, and methods developed to test for associations between species in communities can be used to study serotype distributions among hosts.

African horse sickness (AHS) is a midge-transmitted *Orbivirus* infecting several equine species, with nine serologically distinct types. The virus is endemic in the zebra (*Equus burchelli*) population in the Kruger National Park, South Africa, where all nine serotypes are known to occur [9]. The biting midge *Culicoides imicola* is considered the most significant vector, although other species of *Culicoides* may also be involved [3, 10]. Here, we present further analysis of previously published data [9], to investigate the independence of serotype transmission and distribution and the effect of multiple serotypes on calculations of R_0 for AHS.

METHODS AND MATERIALS

Serum collection

Serum was collected from free-living zebra foals in the Kruger National Park and analysed as described by Barnard [9]. Briefly, between August 1991 and May 1992, foals were captured, age estimated based on tooth eruption and wear, and serum collected. Animals were assigned to monthly age classes except for 7 and 8 months, which could not be distinguished and so were combined into one class. Agar gel immunodiffusion (GP) was used to test for antibodies to AHS, and microneutralization was used to determine serotype-specific antibody titres (NA). Any titre greater than zero was considered positive. There are cross-reacting pairs of serotypes (1 and 2, 3 and 7, 5 and 8, and 6 and 9 [3]); if the specific titre was > 320 for one member of a cross-reacting pair and < 20 in the other, the lower titre was assumed to be the result of cross-reactions and the animal considered negative for that serotype. If both titres were high or low, the animal was considered positive for both.

Because cross-reactions between serotypes may influence the statistical tests used, the analyses were repeated with cross-reacting serotypes combined. If an animal was positive for either of the two members of a cross-reacting pair, it was considered positive for the combination (e.g. if positive for either serotype 1 or 2, it was positive for the combination 1–2). This resulted in five combinations: 1–2, 3–7, 4, 5–8 and 6–9. This will be referred to as the analysis of combined serotypes.

Table 1. Results of Schluter's W^* tests for independent distributions of serotypes within each age class.

Age class	Sample size†	All serotypes	Combined serotypes
5	11	76.3‡	16.1
6	36	87.1‡	65.2‡
7–8	30	117.0‡	74.3‡
9	19	24.2	19.0

* Schluter's W distributed as $\chi^2_{z/2, z}$ where z is sample size.

† Number of zebra in age class.

‡ $P < 0.005$.

Barnard [9] repeatedly sampled four captive zebra foals and determined that they lost maternal antibodies to AHS at 4–6 months of age. There were no apparent differences between serotypes, although the sample size is inadequate to examine this in any detail. In the data analysed here, foals aged 5 months had a bimodal distribution of serotypes, most having positive titres to 3 or fewer serotypes and 4 foals having positive titres to 8 or more serotypes. To prevent maternal antibodies from confounding the analysis, all foals under 5 months of age and the 4 goals aged 5 months which tested positive to 8 or more serotypes were excluded from the analysis. Final sample sizes for each age class are given in Table 1.

Associations

If the distribution of a serotype among hosts is independent of all other serotypes, there should be no covariance between the serotypes. Schluter's variance test [11, 12] compares the observed variance in the number of serotypes per zebra with that expected under the hypothesis of no covariance; if this ratio is significantly different than one there is covariance between the serotypes and the distributions are not independent. This test is equivalent to the summed binomial [13] and Cochran's Q [12], and can be interpreted as comparing the observed distribution of the number of serotypes per zebra with that expected from the sum of independent binomial distributions.

For age classes where there was covariance between the serotypes, pairwise associations were tested using Fisher's exact test (2-tailed; FREQ procedure, SAS [14]). The significance level was adjusted for multiple tests using the Bonferroni correction [15]. All tests for association were carried out on each age class separately.

Generalized linear models (GLM) were used to test for seasonal influences on serotype acquisition, with the number of serotypes positive for each zebra as the dependent variable and age and month of capture as main effects. In addition, the three youngest age classes were tested using only the month of capture as the main effect. The SAS GLM procedure was used for all analyses.

Basic reproduction number

R_0 provides an estimate of the transmission potential of a disease and, in an endemic situation, can be calculated using age-prevalence data [8]. It was assumed that zebra acquire infection with serotype i with an age-independent constant force of infection, λ_i . The proportion infected at age a , $p_i(a)$, is therefore,

$$p_i(a) = 1 - e^{-\lambda_i(a-M)}$$

where M is the age at which maternal antibodies become undetectable. Using maximum likelihood methods [16], λ_i can then be estimated as the value which maximizes the log likelihood,

$$\sum_a -x_{ia} \lambda_i(a-M) + y_{ia} \ln(1 - e^{-\lambda_i(a-M)})$$

where x_{ia} is the number seronegative for serotype i at age a and y_{ia} the number seropositive at age a . As M is not known precisely, it was chosen to maximize the sum (over the nine serotypes) of the log likelihood. The variance of λ can be estimated as

$$\text{var}(\lambda_i) = - \left[\sum_a -y_{ia}(a-M)^2 \left[\frac{e^{-\lambda_i(a-M)}}{1 - e^{-\lambda_i(a-M)}} + \frac{e^{-2\lambda_i(a-M)}}{(1 - e^{-\lambda_i(a-M)})^2} \right] \right]^{-1}$$

The goodness of fit of λ_i for each serotype can be tested by likelihood ratios, comparing the maximum log likelihood (ML) with the likelihood from the fully saturated model (SM), calculated using observed values; $2*(SM - ML) \sim \chi^2_{2, n-3}$ where n is the number of age classes. A common λ was fitted using the data for all nine serotypes, and tested by likelihood ratios (as above) to the fit obtained with individual values of λ_i for each serotype. A significant lack of fit in this test indicates that the nine λ_i s are different. Confidence intervals were also calculated for pairwise comparisons.

R_0 for each serotype is then calculated by

$$R_0 = \lambda_i(L - M)$$

where L is the average lifespan of the host. This method assumes that the serotypes are independently transmitted, that immunity to serotypes is life long (which is thought to be the case for AHS [3]), and that virus prevalence is in equilibrium. It also assumes that the host population is static (i.e. is not growing or decreasing and has a stable age structure). It is reasonable to do this since a zebra population slightly south of those sampled here was reasonably stable from 1984–92 [17]. In an earlier study, Mills and Shenk [18] estimated mortality rates of zebra in the southern population, to give an estimate for L of 72 months. However, they assumed that all animals were killed by lions. To account for other causes of death a value for L of 60 months was used. R_0 was calculated for each serotype from the individual NA tests, and for AHS overall from the GP and summed NA tests (i.e. an animal was considered positive for AHS if positive in any of the nine NA tests).

RESULTS

The overall pattern of prevalence is shown in Fig. 1a, and for each serotype in Fig. 1b. By 9–12 months old, the prevalence of all serotypes approached 100%.

Associations

The distributions of the number of serotypes per zebra in each age class are shown in Fig. 2. There was significant covariance in the 6- and 7–8-month-old age classes both with all serotypes and with serotypes combined, indicating that the serotypes were not independently distributed (Table 1). In the 5 month old age class, there was only significant covariance when all serotypes were considered.

Accordingly, pairwise tests were carried out for the 6- and 7–8-month age classes. Only positive associations were found, i.e. an excess of individuals either positive or negative to both serotypes concerned. In 6-month-old zebra, the only significant pairwise association was between serotypes 6 and 9 (Table 2). This association (by definition) was not found in the analysis with cross-reacting serotypes combined. There were four significant pairwise associations in 7–8-month-old animals: serotypes 4 and 8, 4 and 9, 6 and 9, and 8 and 9 (Table 2). Again, by definition the association between serotypes 6 and 9 was not found in the combined analysis. The associations between serotypes 4 and 8 and 8 and 9 remained significant (4,

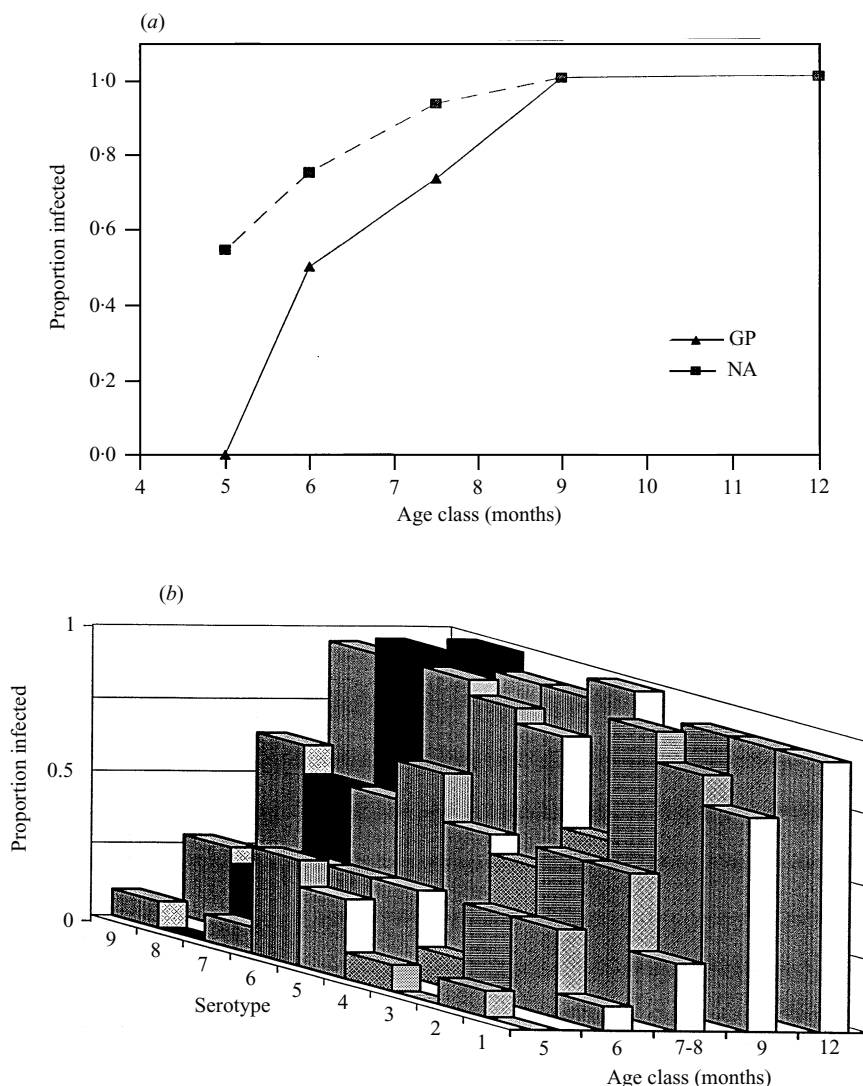


Fig. 1. (a) Overall prevalence of AHS by age. GP: based on prevalence of group-specific antibodies to AHS. NA: based on prevalence of neutralizing antibodies to any serotype. (b) Prevalence of each serotype by age.

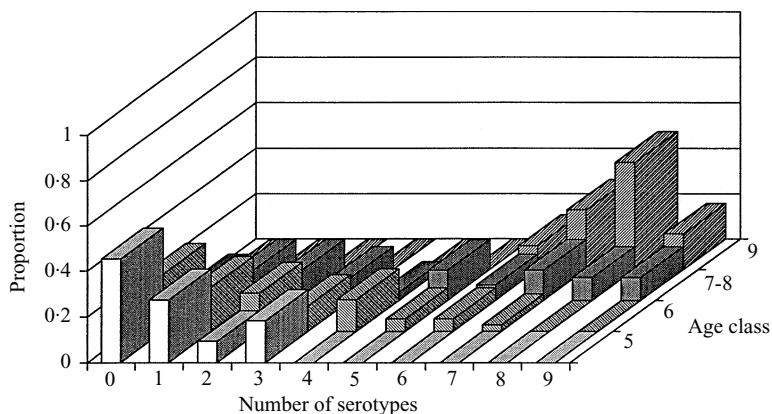


Fig. 2. The distribution of the number of serotypes per zebra by age.

5–8: $P = 0.00015$; 5–8, 6–9: $P = 0.00017$; with the Bonferroni correction an association is significant at the 0.05 level if $P < 0.00125$), while the association

between serotypes 4 and 9 was close to significance (4, 6–9: $P = 0.00131$) in the analysis of combined serotypes.

Table 2. P values for pairwise associations in 6 (above diagonal) and 7–8 (below diagonal)-month-old zebra

	Serotype								
	1	2	3	4	5	6	7	8	9
Serotype 1		0.023	0.538	1	0.22	0.22	1	0.51	1
2	0.11		0.7	1	0.06	0.0007	0.056	0.68	0.004
3	0.11	0.001		1	0.25	1	0.056	1	1
4	0.04	0.29	0.72		0.63	1	0.63	0.31	0.12
5	0.46	0.03	0.48	0.003		0.25	0.25	0.004	0.22
6	0.68	0.14	0.14	0.009	0.02		0.25	0.68	0.0002*
7	0.04	0.01	0.07	0.009	0.16	0.006		0.68	0.22
8	0.46	0.16	0.48	0.00018*	0.001	0.001	0.004		0.18
9	0.43	0.66	0.27	0.00015*	0.007	0.00003*	0.002	0.0004*	

* Significantly different from expected under the hypothesis of no association at the $\alpha = 0.05$ level ($P < 0.0005$, with the Bonferroni correction for multiple tests).

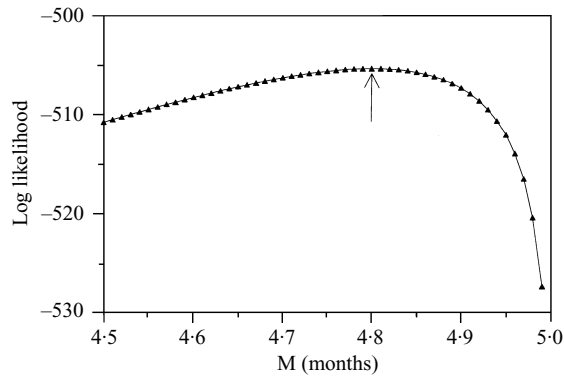


Fig. 3. Sensitivity to M , duration of maternal antibodies, of the sum over serotypes of the log likelihoods. The arrow indicates the maximum value.

Age and month of capture were not correlated ($r = 0.09$, $P = 0.35$), and so could be considered as separate predictors in a GLM. Both age and month of capture were significant factors in the number of serotypes per zebra (age: $F_{1,109} = 113.7$, $P = 0.0001$; month: $F_{1,109} = 9.97$, $P = 0.002$) when all serotypes were considered. With cross-reacting serotypes combined, only age was significant (age: $F_{1,109} = 75.17$, $P = 0.0001$; month: $F_{1,109} = 3.29$, $P = 0.0724$). In the age classes separately, month of capture was only significant in the 6-month class both with all serotypes ($F_{1,35} = 17.24$, $P = 0.004$) and with cross-reacting serotypes combined ($F_{1,35} = 17.24$, $P = 0.0002$).

Basic reproduction number

The summed log likelihood over the nine serotypes was maximized at $M = 4.8$ months (confidence interval 4.65–4.9, Fig. 3). Using this value, χ^2 tests

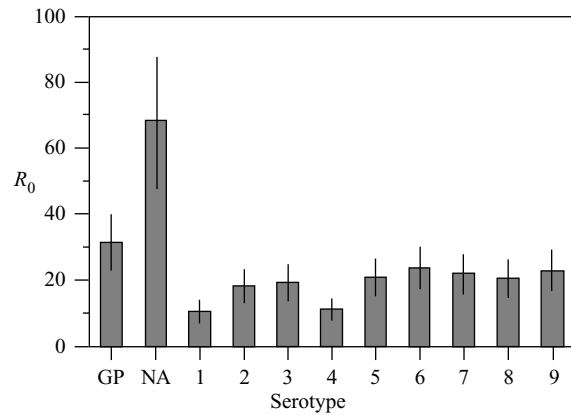


Fig. 4. R_0 (mean \pm 95% C.I.) for AHS overall and for each of the nine serotypes individually, calculated using $M = 4.8$. GP, NA: as in Fig. 1. S1–S9; based on prevalence of neutralizing antibodies to each serotype alone.

indicate that estimated values of λ_i adequately fit the data (likelihood ratios < 16.5 , $P > 0.01$; with the Bonferroni correction $P < 0.005$ is required for a significant lack of fit at the 0.05 level). Estimates of λ_i for each serotype were significantly different (likelihood ratio = 32.6, $P < 0.005$). Comparing confidence intervals indicates that λ_1 and λ_4 were significantly smaller than several other serotypes ($\lambda_1 < \lambda_5, \lambda_6, \lambda_7, \lambda_8, \lambda_9$; $\lambda_4 < \lambda_5, \lambda_6, \lambda_7, \lambda_9$). R_0 for individual serotypes varied from 10.5 (serotype 1) to 23.4 (serotype 6, Fig. 4). As R_0 is a multiple of λ_i , the same significance relationships apply; estimates with overlapping confidence intervals in Figure 4 are not significantly different. Looking at AHS as a whole, the GP method gives a much lower estimate of $R_0 = 31$ than the combined NA tests ($R_0 = 68$, Fig. 4); the confidence intervals do not overlap (Fig. 4).

DISCUSSION

The serotypes of AHS were not distributed independently in this zebra population, but it is difficult to determine the reason. Associations between serotypes of viruses can arise in several ways. Infection with one serotype may predispose an animal to infection with another serotype or co-occurring serotypes may be more easily transmitted (facilitation); cross-reactive immunity between serotypes may cause negative associations; or heterogeneity in host-vector transmission rates may create non-random distributions of serotypes. Any of these may lead to non-independent distributions of serotypes among hosts (within an age class), but the last does not require non-independent transmission, i.e. there are no biological interactions between serotypes.

Facilitation would be expected to lead to positive associations between specific serotypes, and can act in two ways, either sequentially or simultaneously. In sequential facilitation, infection with one parasite increases susceptibility to another species or serotype of parasite. If this was the case, when individuals are repeatedly sampled over time one serotype should consistently appear first. In cohort data, as presented here, the prevalence of one of the associated serotypes alone should be the higher than the combination of serotypes, while the prevalence of single infections with the other serotype (the one which requires the facilitation) should be the lowest. No such patterns are apparent in these data, as there were relatively few single infections with any of the three serotypes which had pairwise associations (serotypes 4, 8 and 9). With simultaneous facilitation, a transmission event is more likely to result in infection if more than one serotype is present. This would result in the pattern observed in the 7–8-month-old zebra, where some serotypes are present primarily together and rarely or never alone. Further information on the course of infection in animals simultaneously infected with more than one serotype is necessary to determine if facilitation is responsible for the associations observed.

Cross-immunity occurs between AHS serotypes (see above), and would be expected to lead to negative associations between serotypes. However, if serotype identification is compromised by cross-immunity (as for AHS), then artefactual positive associations may result. Tests with cross-reacting serotypes combined are more conservative, and will reduce this problem. Here, since these tests generally showed the same patterns as the full data set, the non-random patterns demonstrated appear not to be a result of cross-

immunity. This comparison should be made in systems where there is cross-immunity between serotypes, to eliminate cross-reactions as the source of associations. Clearly, it is important that patterns of cross-immunity are carefully considered when interpreting the epidemiology of serotypes.

Low levels of heterogeneity in transmission rates could lead to non-random distributions of serotypes among zebra. Several types of heterogeneity in virus transmission between hosts and vectors may exist, such as heterogeneity in biting rates (e.g. some zebra may be bitten more often than others), spatial or temporal variations in the distribution of hosts or vectors, and differences in susceptibility to different serotypes (e.g. genetic differences in susceptibility of *Culicoides variipennis* to serotypes of bluetongue virus [19, 20]).

It cannot be determined from these results whether the observed associations are stable; that is, would the same associations be observed in other years or in other populations. If the associations are due to facilitation, cross-immunity or genetic differences in susceptibility, it would be expected that the same associations would be seen in other years and possibly in other populations. However, if the associations were due to other types of heterogeneity, while non-independent distributions would be expected, the same associations would not necessarily be expected in other years or populations.

Pairwise associations between serotypes were only present in the 7–8-month-old foals. These animals are very susceptible to infection, as maternal antibodies will have waned to undetectable levels. Older foals are infected with most serotypes, and so associations would be difficult to detect, while younger animals may not have not acquired sufficient infections for associations to be significant. The lack of similar associations in the 6-month-old zebra may reflect the small sample sizes, low statistical power of the tests (resulting from the large number of pairwise tests needed), or the differences in the distributions of month captured between the 7–8- and 6-month-old zebra [9]. The pairwise associations in the 7–8-month-old foals appear to result from more animals infected with all three (4, 8 and 9) serotypes than would be expected, although it was impractical to test 3-way associations due to low expected values.

These values of R_0 are among the highest reported for veterinary diseases [21]. The large population and year-round addition of new susceptible animals to the population allow the virus to persist with high levels of transmission. Although the method used here to

calculate R_0 for the individual serotypes assumes independent transmission, the relative differences are still informative, even if there are some interactions between the serotypes. The variation in R_0 by serotype indicates that there are likely to be differences in infectivity of different serotypes for midges of zebra, or both.

It is interesting to note that serotype 9, considered avirulent in horses in South Africa [3], has a higher R_0 than other serotypes. Conversely, serotype 4, the serotype involved in recent epidemics in the Iberian peninsula, has a lower R_0 . There are variations in virulence within serotypes [22], and the relationship between serotypes, geographical distribution, virulence and transmission are poorly understood. In addition, virulence is usually determined by the impact of the virus on horses, and it is not known if this is relevant to infections in zebra. Estimates of R_0 are specific to combinations of host, vector and place, and hence further research on the basic reproduction number for serotypes of AHS under different conditions is needed.

In general, it is difficult to establish independent transmission from epidemiological data since non-independent distributions of serotypes among hosts need not imply non-independent transmission. An upper limit on the estimate of R_0 which does not depend on assumptions of independent transmission can be derived from the prevalence of exposure to any serotype of AHS (NA in Figs. 1 and 4). If, however, the assumption of independent transmission is made, a reliable lower limit on the estimate is the highest individual R_0 , as any control measures designed using this estimate would exceed the levels required for the other serotypes. This is preferable to the use of a weighted average [7], which may underestimate the level of control needed for the most transmissible serotypes. For AHS, the highest individual value of R_0 was 23, whereas if all serotypes are considered together the upper estimate of R_0 was 68. This is a substantial difference and would have a significant effect on the formulation of a control program. It is important, therefore, to establish the extent to which AHS serotypes may be regarded as being independently transmitted.

ACKNOWLEDGEMENTS

We thank Dan Haydon and Sunetra Gupta for discussions of strain variation. We also thank Joanne Webster for helpful discussions. The assistance of Christl Donnelly on the maximum likelihood methods

is gratefully acknowledged. C.C.L. was supported by a BBSRC Link Grant and M.E.J.W. by the Royal Society.

REFERENCES

1. Forsyth KP, Philip G, Smith T, Kum E, Southwell B, Brown GV. Diversity of antigens expressed on the surface of erythrocytes infected with mature *Plasmodium falciparum* parasites in Papua New Guinea. *Am J Trop Med Hyg* 1989; **41**: 259–65.
2. Woolhouse MEJ, McNamara JJ, Hargrove JW, Bealby KA. Distribution and abundance of trypanosome (subgenus *Nannomonas*) infections of the tsetse fly *Glossina pallidipes* in southern Africa. *Molec Ecol* 1996; **5**: 11–8.
3. Coetzer JAW, Erasmus BJ. African horse sickness. In: Coetzer JAW, Thomson GR, Tustin RC, eds. *Infectious diseases of livestock with special reference to southern Africa*. Oxford: Oxford University Press, 1994: 460–75.
4. Verwoerd DW, Erasmus BJ. Bluetongue. In: Coetzer JAW, Thomson GR, Tustin RC, eds. *Infectious diseases of livestock with special reference to southern Africa*. Oxford: Oxford University Press, 1994: 443–59.
5. Domingo E, Mateu MG, Martinez MA, Dopazo J, Moya A, Sobrino F. Genetic variability and antigenic diversity of foot-and-mouth disease virus. In: Kurstak E, Marusyk RG, Murphy FA, van Regenmortel MHV, eds. *Applied virology research*, vol. 2. New York: Plenum, 1990: 233–66.
6. Kettle DS. *Medical and veterinary entomology*. Wallingford: CAB International, 1990: 455–7.
7. Gupta S, Trenholme K, Anderson RM, Day KP. Antigenic diversity and the transmission dynamics of *Plasmodium falciparum*. *Science* 1994; **263**: 961–3.
8. Anderson RM, May RM. *Infectious diseases of humans: dynamics and control*. Oxford: Oxford University Press, 1991.
9. Barnard, BJH. Circulation of African horsesickness virus in zebra (*Equus burchelli*) in the Kruger National Park, South Africa, as measured by the prevalence of type specific antibodies. *Onderstepoort J Vet Res* 1993; **60**: 111–7.
10. Meiswinkel R, Nevill EM, Venter GJ. Vectors: *Culicoides* spp. In: Coetzer JAW, Thomson GR, Tustin RC, eds. *Infectious diseases of livestock with special reference to southern Africa*. Oxford: Oxford University Press, 1994: 68–89.
11. Schluter D. A variance test for detecting species associations, with some example applications. *Ecology* 1984; **65**: 998–1005.
12. McCulloch CE. Variance tests for species associations. *Ecology* 1985; **65**: 1676–81.
13. Lotz JM, Font WF. The role of positive and negative interspecific associations in the organization of communities of intestinal helminths of bats. *Parasitol* 1991; **103**: 127–38.
14. SAS. *SAS User's guide*. Cary, NC: SAS Institute, 1988.
15. Rice WR. Analyzing tables of statistical tests. *Evolution* 1989; **43**: 223–5.

16. Williams BG, Dye C. Maximum likelihood for parasitologists. *Parasitol Today* 1994; **10**: 489–93.
17. Mills MGL, Biggs HC, Whyte IJ. The relationship between rainfall, lion predation and population trends in African herbivores. *Wildl Res* 1995; **22**: 75–88.
18. Mills MGL, Shenk TM. Predator–prey relationships: the impact of lion predation on wildebeest and zebra populations. *J Anim Ecol* 1992; **61**: 693–702.
19. Tabachnick, WJ. Genetic control of oral susceptibility to infection of *Culicoides variipennis* with bluetongue virus. *Am J Trop Med Hyg* 1991; **45**: 666–71.
20. Tabachnick, WJ. Microgeographic and temporal genetic variation in populations of the bluetongue virus vector *Culicoides variipennis* (Diptera: Ceratopogonidae). *J Med Entomol* 1992; **29**: 384–94.
21. Woolhouse MEJ, Bundy DAP. Epidemiological aspects of vaccination programs. In: Pastoret PP, Blancou J, Vannier P, Verschueren C, eds. *Veterinary vaccinology*. London: Elsevier Press. In press.
22. Laegreid WW, Skowronek A, Stone-Marschat M, Burrage T. Characterization of virulence variants of African horse sickness virus. *Virology* 1993; **135**: 836–9.