

Stage-structured transmission of phocine distemper virus in the Dutch 2002 outbreak

Petra Klepac^{1,2,†}, Laura W. Pomeroy^{3,*},†, Ottar N. Bjørnstad^{1,2,4}, Thijs Kuiken⁵,
Albert D. M. E. Osterhaus⁵ and Jolianne M. Rijks^{5,6}

¹Department of Biology, ²Center for Infectious Disease Dynamics, and ⁴Department of Entomology, The Pennsylvania State University, University Park, PA 16802, USA

³Department of Bioinformatics, College of Medicine, The Ohio State University, Columbus, OH 43210, USA

⁵Department of Virology, Erasmus Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands

⁶Faculty of Veterinary Medicine, Dutch Wildlife Health Center, Utrecht University, Yalelaan 1, 3584 CL Utrecht, The Netherlands

Heterogeneities in transmission among hosts can be very important in shaping infectious disease dynamics. In mammals with strong social organization, such heterogeneities are often structured by functional stage: juveniles, subadults and adults. We investigate the importance of such stage-related heterogeneities in shaping the 2002 phocine distemper virus (PDV) outbreak in the Dutch Wadden Sea, when more than 40 per cent of the harbour seals were killed. We do this by comparing the statistical fit of a hierarchy of models with varying transmission complexity: homogeneous versus heterogeneous mixing and density-versus frequency-dependent transmission. We use the stranding data as a proxy for incidence and use Poisson likelihoods to estimate the ‘who acquires infection from whom’ (WAIFW) matrix. Statistically, the model with strong heterogeneous mixing and density-dependent transmission was found to best describe the transmission dynamics. However, patterns of incidence support a model of frequency-dependent transmission among adults and juveniles. Based on the maximum-likelihood WAIFW matrix estimates, we use the next-generation formalism to calculate an R_0 between 2 and 2.5 for the Dutch 2002 PDV epidemic.

Keywords: phocine distemper virus; harbour seals; *Phoca vitulina*; stage-structured transmission; who acquires infection from whom matrix; next-generation matrix

1. INTRODUCTION

Heterogeneities in pathogen transmission can change invasion criteria (Woolhouse *et al.* 1997; Newman 2005; Ferrari *et al.* 2006) and enhance spatial spread through superspreading events (Lloyd-Smith *et al.* 2005). Heterogeneities in transmission arise due to underlying gender, age-specific, behavioural, immunological and genetic variations in susceptibility and infectiousness (Anderson & May 1991; Altizer *et al.* 2003; Perkins *et al.* 2003; Lloyd-Smith *et al.* 2005). The most common theoretical approach to account for such heterogeneities is the ‘who acquires infection from whom’ (WAIFW) matrix (Anderson & May 1984, 1985, 1991; Schenzle 1984; Dobson 2004), which recognizes a refined transmission rate, $\beta_{i,j}$, at which an infectious individual of class j will infect a susceptible individual of class i (Anderson & May 1991). The separate classes can encompass gender, age, stage, social, immunological, physiological or behavioural differences.

While the WAIFW matrix has proven to be of great theoretical utility (Schenzle 1984; Anderson & May 1985; Dobson 2004; Kanaan & Farrington 2005), empirical approaches to its estimation and characterization have often proven difficult because of the lack of relevant data. Various efforts have employed contact tracing

(Ferguson *et al.* 2001; Fraser *et al.* 2004) or inferred contacts (Edmunds *et al.* 1997; Huang & Rohani 2006; Wallinga *et al.* 2006) to determine how a pathogen is transmitted among different classes within the host population. In this paper, we investigate whether it is possible to estimate the elements in the WAIFW matrix from detailed age- or stage-structured incidence data. We also use our statistical approach to compare density versus frequency dependence as the most parsimonious model for transmission. We ask four nested questions regarding how to best model the transmission of phocine distemper virus (PDV) among Dutch harbour seals during the 2002 epidemic: (i) is there evidence of stage-structured transmission within the host population? (ii) Given stage structure in transmission, how well can we identify the WAIFW elements by applying maximum-likelihood estimation to the time series of stage-specific incidence? (iii) Is a model of density-dependent or frequency-dependent transmission best supported by the incidence data? Finally, (iv) how can we combine our maximum-likelihood method for estimating the WAIFW matrix with the theoretical next-generation formalism to estimate R_0 for PDV?

To address these questions, we investigate the spread of PDV in harbour seals (*Phoca vitulina*) in The Netherlands during the 2002 epidemic (Rijks *et al.* 2005). PDV is a single-stranded, negative sense RNA virus that is a member of the morbillivirus genus, family *Paramyxoviridae* (Cosby *et al.* 1988; Mahy *et al.* 1988; Osterhaus & Vedder 1988). In each individual, disease typically spans a two-week

* Author and address for correspondence: 220 Biomedical Research Tower, 460 West 12th Avenue, Columbus, OH 43210, USA (pomeroy.26@osu.edu).

† These authors contributed equally to this work.

period, including both the latent and infectious disease stages (Osterhaus *et al.* 1989; Harder *et al.* 1990; Baker 1992; Grenfell *et al.* 1992).

Two outbreaks of PDV have affected seal populations throughout the entire North Sea region: the first outbreak occurred in 1988, in which 18 000–23 000 harbour seals died (Hall *et al.* 2006; Härkönen *et al.* 2006). This mass mortality event caused by the viral epidemic began on the Danish island of Ånholt on 12 April 1988 and ended within the calendar year (Dietz *et al.* 1989; Hall *et al.* 2006; Härkönen *et al.* 2006). A second PDV outbreak occurred in 2002, also starting at the island of Anholt: initial cases of harbour seal stranding and mortality occurred on 4 May 2002. This time, somewhere between 22 000 and 30 000 harbour seals died, resulting in the largest recorded mass mortality event in marine mammals (Jensen *et al.* 2002; Hall *et al.* 2006; Härkönen *et al.* 2006).

During the second epidemic in The Netherlands, the first case of PDV was found on 16 June 2002 on Vlieland and the local epidemic ceased at the end of November, as fully described in Rijks *et al.* (2005). In that time period, 2284 seals were stranded along the Dutch coast, including 2279 harbour seals and 5 grey seals. The index case was a subadult; moreover, the median stranding date of all subadults was significantly earlier than the median stranding date of both juveniles and adults (Rijks *et al.* 2005). Together, the available stranding data suggest a possible role of stage-structured disease transmission and heterogeneous host mixing in this 2002 Dutch epidemic. While previous models describing the spread of PDV throughout the North Sea have assumed homogeneous mixing among different harbour seal age or stage classes (Grenfell *et al.* 1992; De Koeijer *et al.* 1998; Swinton 1998; Swinton *et al.* 1998) and have focused on disease in the North Sea harbour seal metapopulation, we have a unique opportunity to use a detailed, stage-structured time series of stranded PDV cases to estimate heterogeneous transmission in a population of harbour seals in this paper. In addition to explicitly studying transmission dynamics among harbour seals, our work is novel in that it introduces a statistical framework for testing for non-homogeneous mixing from stage-structured incidence data.

2. MATERIAL AND METHODS

Stranded seals were classified into stages based on the body length of carcasses: relationships between age and body lengths were determined by precise ageing in approximately 25 per cent of seal carcasses by counting the cementum layers of a canine tooth, determining the body lengths for the precisely aged seals, and extrapolating these to the entire dataset. The juvenile class contained female seals less than 90 cm and male seals less than 95 cm. Subadults included females with body lengths between 90 and 120 cm and males with body lengths between 95 and 130 cm. Lastly, the adult category contained female seals with body lengths greater than 120 cm and males with body lengths greater than 130 cm. Using these body length classifications, the juvenile class contained most of the pups of the year, while the subadult class included most of the 1- and 2-year-old females and 1- to 3-year-old males. Finally, the adult class included most of the females older than 2 years and males older than 3 years.

The epidemic dynamics were captured with a susceptible–infected–removed (SIR) model, dividing the population into

three categories based on their epidemiological state. Susceptible individuals have never been infected nor exposed to the virus. Infected individuals harbour the virus and are able to transmit the infection to susceptible individuals. Lastly, removed individuals have previously been infected and either recover from the disease with conferred lifelong immunity or are removed from the system due to disease-induced mortality. Our model is defined in discrete time, with each time step equal to 1 day.

In addition to these three epidemic classes (susceptible, infected and removed), we divide the seal population into three demographic classes—juvenile, subadult and adult—resulting in a model with nine total categories (equation (2.1)).

$$N = \begin{pmatrix} S & I & R \\ n_{1,1} & n_{1,2} & n_{1,3} \\ n_{2,1} & n_{2,2} & n_{2,3} \\ n_{3,1} & n_{3,2} & n_{3,3} \end{pmatrix} \begin{matrix} \text{juveniles} \\ \text{subadults} \\ \text{adults} \end{matrix} \tag{2.1}$$

The initial population size (N_0) was fixed at 5400, based on pre-epizootic population censuses (Ries *et al.* 1998; Trilateral Seal Expert Group 2001); at the start of the epidemic, all 5400 seals were assumed to be susceptible so that $N_0 = S_0 = 5400$. For each element $n_{i,j}$ in the matrix, the subscript i designates the stage structure (1, juveniles; 2, subadults; 3, adults), and the subscript j designates the epidemic category (1, susceptible individuals; 2, infected individuals; 3, recovered or dead individuals). For example, $n_{3,2}$ is the total number of infected adults. The matrix N (equation (2.1)) can be organized into a population vector by stacking the rows of the matrix, as in equation (2.2), where $'$ specifies a vector transpose. The population vector then consists of all juvenile classes (susceptible, infected and removed juveniles), followed by all subadults, and then all adults.

$$\mathbf{n} = (n_{1,1} \ n_{1,2} \ n_{1,3} \ | \ n_{2,1} \ n_{2,2} \ n_{2,3} \ | \ n_{3,1} \ n_{3,2} \ n_{3,3})' \tag{2.2}$$

Let $\beta_{i,j}$ be the probability per time step of disease transmission between a susceptible individual in demographic class i and an infected individual in demographic class j , where the time step is 1 day. The transmission, or WAIFW matrix, is then a 3 by 3 matrix, $\beta = \beta_{i,j}$. We considered four different transmission scenarios and built four different models. For the first model, we considered homogeneous mixing in different stage classes and equal transmission rates ($\beta_{i,j} = \beta$ for all i,j). For the second model, we assumed weak heterogeneous mixing among the host population: within-stage transmission—the diagonal of the transmission matrix—was allowed to differ from between-stage transmission, which was designated by the off-diagonal elements. This difference was scaled by a coefficient, k . The transmission matrix is then

$$\beta = \begin{pmatrix} k\beta & \beta & \beta \\ \beta & k\beta & \beta \\ \beta & \beta & k\beta \end{pmatrix} \tag{2.3}$$

For the third model, we assumed strong heterogeneous mixing between stages by allowing the transmission term to vary with the interactions within and between stages (equation (2.4)).

$$\beta = \begin{pmatrix} \beta_{1,1} & \beta_{1,2} & \beta_{1,3} \\ \beta_{2,1} & \beta_{2,2} & \beta_{2,3} \\ \beta_{3,1} & \beta_{3,2} & \beta_{3,3} \end{pmatrix} \tag{2.4}$$

The transmission matrix (equation (2.4)) is assumed to be symmetrical: $\beta_{i,j} = \beta_{j,i}$. This property captures the assumption that the probability of the two hosts in different stages infecting each other is equal. In other words, differences in susceptibility or infectiousness among stage classes do not significantly impact on transmission events (Anderson & May 1991). It would be possible to change this assumption, but our model is already quite parameter-rich and additional parameters would lead to further identifiability problems (see below).

In this framework, the total force of infection for each class, ϕ_i , or the probability per unit time for a susceptible individual in a certain demographic class to become infected (Diekmann & Heesterbeek 2000), includes all the possible transmission events with infected cases in all demographic classes,

$$\phi_i(\mathbf{n}(t)) = 1 - \exp\left(-\sum_j \beta_{i,j} n_{j,2}(t)\right). \tag{2.5}$$

Once infected, individuals of all stages remain infected for an average of 14 days ($\gamma = 1/14$ days). The epidemic transitions matrix for a stage class i is then

$$\mathbf{A}_i = \begin{pmatrix} 1 - \phi_i(\mathbf{n}(t)) & 0 & 0 \\ \phi_i(\mathbf{n}(t)) & 1 - \gamma & 0 \\ 0 & \gamma & 1 \end{pmatrix}. \tag{2.6}$$

This model (equation (2.6)) captures the density-dependent contact process, where force of infection (equation (2.5)) is proportional to the number of infected individuals. In this case, the number of contacts an individual makes in a unit of time is proportional to the total population size.

When the number of contacts per infected individual per unit time is constant, the process of transmission is known as frequency-dependent transmission. In this case, the force of infection is proportional to the proportion of infected individuals in the population, or I/N ,

$$\phi_i(\mathbf{n}(t)) = 1 - \exp\left(-\sum_j \frac{\beta_{i,j} n_{j,2}(t)}{N_t}\right), \tag{2.7}$$

where N_t is the total living population at time t . The epidemic transitions for stage i are then simply given by \mathbf{A}_i (equation (2.6)), where ϕ_i is defined in equation (2.7) and γ in the (3, 2) location of the matrix is multiplied by the recovery rate r .

Each of the demographic classes—juveniles, subadults and adults—has a respective epidemic transition matrix, \mathbf{A}_1 , \mathbf{A}_2 and \mathbf{A}_3 . Epidemic transitions for the entire population vector \mathbf{n} are given by the block-diagonal matrix $\mathbf{A}(\mathbf{n})$, with matrices \mathbf{A}_1 , \mathbf{A}_2 and \mathbf{A}_3 on the diagonal, and zeros elsewhere (equation (2.8)).

$$\mathbf{A}(\mathbf{n}) = \begin{pmatrix} \mathbf{A}_1 & 0 & 0 \\ 0 & \mathbf{A}_2 & 0 \\ 0 & 0 & \mathbf{A}_3 \end{pmatrix}. \tag{2.8}$$

Since we assumed that the epidemic dynamics are fast relative to demography, there are no transitions between stage classes, and the remaining elements of the block-diagonal matrix $\mathbf{A}(\mathbf{n})$ are zero. The epidemic trajectory is given by multiplying the population vector at time t , $\mathbf{n}(t)$, with the transition matrix $\mathbf{A}(\mathbf{n})$

$$\mathbf{n}(t + 1) = \mathbf{A}[\mathbf{n}(t)]\mathbf{n}(t). \tag{2.9}$$

Due to the underlying stage structure, there are three infected classes in this model. For models with multiple

classes, R_0 can be derived using the next-generation method (Diekmann et al. 1990; de Jong et al. 1994; Diekmann & Heesterbeek 2000; van den Driessche & Watmough 2002, Allen & van den Driessche 2008), where R_0 is given by the spectral radius, ρ , or the dominant eigenvalue of the next-generation matrix: $\mathbf{F}(\mathbf{I} - \mathbf{T})^{-1}$

$$R_0 = \rho[\mathbf{F}(\mathbf{I} - \mathbf{T})^{-1}]. \tag{2.10}$$

To find the next-generation matrix of a model with s compartments of which r are infected, we let $n = n_1, \dots, n_s$ be the number of individuals in each compartment; \mathbf{F} is the vector of new infections, and \mathbf{T} is the vector of all other transitions, so that $n(t) = \mathbf{F}(t) + \mathbf{T}(t)$. Matrices \mathbf{F} and \mathbf{T} are obtained by differentiation with respect to the infected states and evaluation at the disease-free equilibrium.

$$\mathbf{F} = \left[\frac{\partial F_i}{\partial n_j}(n_0) \right] \quad \text{and} \quad \mathbf{T} = \left[\frac{\partial T_i}{\partial n_j}(n_0) \right], \tag{2.11}$$

where $i, j = 1, \dots, r$ and n_0 is the disease-free equilibrium, at which the population remains in the absence of the disease (van den Driessche & Watmough 2002). The (j, k) entry of $(\mathbf{I} - \mathbf{T})^{-1}$ is the average amount of time an infective individual that was introduced into compartment k spent in compartment j during its lifetime. The (i, j) entry of \mathbf{F} is the rate at which infected individuals in compartment j produce new infections in compartment i . Therefore, the entry (i, k) in the generation matrix $\mathbf{F}(\mathbf{I} - \mathbf{T})^{-1}$ is the expected number of new infections in compartment i produced by an individual originally introduced into compartment k .

The matrix \mathbf{F} shows the influx of new infections to the infectious compartments. Since we assumed that there are no transitions between the infectious classes due to population growth during this acute PDV outbreak, matrix \mathbf{T} reflects the rates at which individuals are leaving the infectious compartments due to recovery or death, accounting for the large number of seal deaths attributable to PDV. At the disease-free equilibrium, the population consists wholly of susceptible individuals, so that

$$\mathbf{n}_0 = [n_{1,1}(0) \ 0 \ 0 \ n_{2,1}(0) \ 0 \ 0 \ n_{3,1}(0) \ 0 \ 0]^T, \tag{2.12}$$

where T again designates vector transpose.

\mathbf{F} and \mathbf{T} were constructed as follows:

$$\mathbf{F} = \begin{pmatrix} \beta_{1,1}n_{1,1}(0) & \beta_{1,2}n_{1,1}(0) & \beta_{1,3}n_{1,1}(0) \\ \beta_{2,1}n_{2,1}(0) & \beta_{2,2}n_{2,1}(0) & \beta_{2,3}n_{2,1}(0) \\ \beta_{3,1}n_{3,1}(0) & \beta_{3,2}n_{2,1}(0) & \beta_{3,3}n_{3,1}(0) \end{pmatrix} \quad \text{and} \tag{2.13}$$

$$\mathbf{T} = \begin{pmatrix} 1 - \gamma & 0 & 0 \\ 0 & 1 - \gamma & 0 \\ 0 & 0 & 1 - \gamma \end{pmatrix}.$$

The next-generation matrix is thus

$$\mathbf{F}(\mathbf{I} - \mathbf{T})^{-1} = \begin{pmatrix} \frac{\beta_{1,1}n_{1,1}(0)}{\gamma} & \frac{\beta_{1,2}n_{1,1}(0)}{\gamma} & \frac{\beta_{1,3}n_{1,1}(0)}{\gamma} \\ \frac{\beta_{2,1}n_{2,1}(0)}{\gamma} & \frac{\beta_{2,2}n_{2,1}(0)}{\gamma} & \frac{\beta_{2,3}n_{2,1}(0)}{\gamma} \\ \frac{\beta_{3,1}n_{3,1}(0)}{\gamma} & \frac{\beta_{3,2}n_{2,1}(0)}{\gamma} & \frac{\beta_{3,3}n_{3,1}(0)}{\gamma} \end{pmatrix}. \tag{2.14}$$

R_0 is given by the dominant eigenvalue of the next-generation matrix (equation (2.14)). We determined R_0 using equations

Table 1. Initial model conditions obtained from the literature. (Total population size (N) was estimated for the entire Dutch harbour seal population, comprising seals that were on land and in the water at any given time.)

parameter	value	reference
N (total population size)	5401	Ries <i>et al.</i> (1998) and Trilateral Seal Expert Group (2001)
infectious period	14 days	Swinton <i>et al.</i> (1998)
gamma (inverse of the infectious period)	1/14 days	Swinton <i>et al.</i> (1998)
population stage structure	15% juveniles 36% subadults 49% adults	Abt (2002)
epidemic stage structure	$S=810$ juveniles + 1944 subadults + 2646 adults $I=1$ subadult $R=0$ individuals	n.a.
length of epidemic	180 days	Rijks <i>et al.</i> (2005)
r	0.4	Klepac (2007)

(2.10) and (2.14) and our estimate of β . We derive and evaluate the expression for R_0 for the frequency-dependent transmission in an analogous way.

Initial model conditions for the total population size (N_0), the infectious period, gamma (the inverse of the infectious period) population stage structure and the length of the epidemic were derived from the literature (table 1). The three nested models—homogeneous mixing with density dependence, weak heterogeneous mixing with density dependence and strong heterogeneous mixing with density dependence—were compared using the likelihood ratio test (LRT). Two models—strong heterogeneous mixing with density dependence and strong heterogeneous mixing with frequency dependence—were compared using Akaike's information criterion (AIC). Subsequently, p -values were calculated to determine which model best fits the data.

We use data on stranded seals as a proxy for incidence. To obtain estimates for the WAIFW matrix, we used maximum-likelihood techniques to find the values of the matrix elements that best fit the stage-specific incidence data (Rijks *et al.* 2005) and the probability of observing a stranded seal, p (equation (2.15)). The probability of observation is a compound variable encompassing both the probability that a given seal, once infected by PDV, will strand and that the stranded seal will be encountered and observed. We assume Poisson likelihoods for disease incidence.

$$y \sim Po(pI_t). \quad (2.15)$$

Data for this model consisted of stranded seals from the Dutch islands of Vlieland, Terschelling, Ameland, Schiermonnikoog and Texel, and from the mainland provinces of Friesland, Groningen and Noord Holland. Point estimates were located by minimizing the negative log likelihood of the data using simulated annealing (Belisle 1992) as implemented by the 'optim' function in R (R Development Core Team 2006).

All model building and parameter estimations were performed using R v. 2.3.1 (R Development Core Team 2006). The next-generation estimates of R_0 were performed using MATHEMATICA v. 6 (Wolfram Research 2007).

3. RESULTS

The data were stratified by stage class (figure 1). The four models created were fitted to the data. The three nested models were compared using the LRT and the models comparing frequency and density dependence

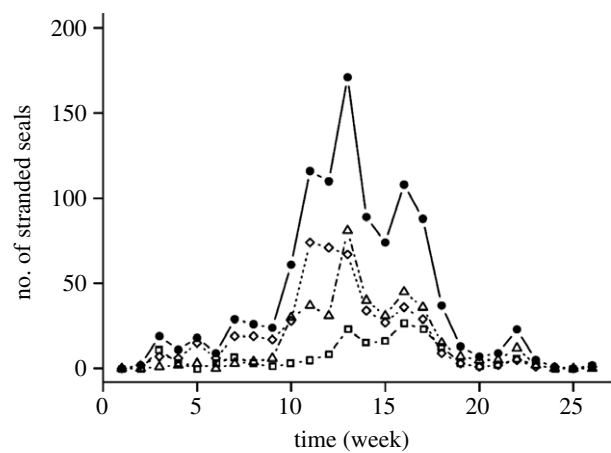


Figure 1. Temporal stranding of harbour seals. The total number of harbour seals stranded and the number of harbour seals stranded stratified by stage class are shown for each week of the PDV epidemic (circles, all seals; squares, juveniles; diamonds, subadults; triangles, adults). The first week began in mid June, while the last cases were counted at the end of November.

were compared using AIC. The first model, homogeneous mixing with density dependence, implies complete lack of stage structure in the population. Results from the model selection tests (table 2) show that the model with slight heterogeneous mixing and density dependence has a better fit to the data than the model with homogeneous mixing and density dependence ($p=0.001$). The strong heterogeneous mixing model with density dependence fits the data better still ($p=0.0007$; table 2). Comparing the strong heterogeneous mixing model with density dependence ($-\log$ likelihood = 248.7909) with the strong heterogeneous mixing model with frequency dependence ($-\log$ likelihood = 255.62) shows that the density-dependent model fits the data better, since the two models are equal in the number of parameters estimated. This results in AIC values that support the strong heterogeneous mixing model with density dependence over the strong heterogeneous mixing model with frequency dependence by 13.8 units. When comparing the set of all four models, the best-fit model overall was the model with the strong heterogeneous mixing, which permitted unique within- and between-stage interactions and density dependence (table 2).

Table 2. Model selection using the LRT. (Three models, used to describe the spread of PDV in the Dutch Wadden Sea, were compared to see which best fit the data. These models include: homogeneous mixing with uniform β , slight heterogeneous mixing with $k\beta$ on the diagonal, and the full stage-structure model with the symmetrical β matrix. The likelihood ratio, degrees of freedom and p -values are listed for each pairwise model comparison. Overall, the model incorporating full stage structure and symmetrical β matrix was the best-fit model to the data.)

models being compared			likelihood ratio	degrees of freedom	p -value	better-fit model
homogeneous mixing with density dependence	v.	slight heterogeneous mixing with density dependence	5.38	1	0.00104	slight heterogeneous mixing with density dependence
slight heterogeneous mixing with density dependence	v.	strong heterogeneous mixing with density dependence	9.63	4	0.000698	strong heterogeneous mixing with density dependence

Table 3. Parameter point estimates in the full stage-structure model from one-dimensional likelihoods. (Point estimates for each element in the symmetrical β matrix in the model incorporating full stage structure were obtained by maximum-likelihood methods. The intra-stage transmission terms are designated by $\beta_{1,1}$ for juveniles, $\beta_{2,2}$ for subadults and $\beta_{3,3}$ for adults. Inter-stage transmission is symmetrical. Interaction between juveniles and subadults is designated by $\beta_{1,2}=\beta_{2,1}$, while transmission between juveniles and adults are designated by $\beta_{1,3}=\beta_{3,1}$. Finally, the transmission between subadults and adults is designated by $\beta_{2,3}=\beta_{3,2}$.)

parameter	demographic classes	value
$\beta_{1,1}$	juvenile–juvenile	3.26×10^{-5}
$\beta_{2,2}$	subadult–subadult	1.93×10^{-5}
$\beta_{3,3}$	adult–adult	6.42×10^{-6}
$\beta_{1,2}=\beta_{2,1}$	juvenile–subadult	8.99×10^{-5}
$\beta_{1,3}=\beta_{3,1}$	juvenile–adult	2.92×10^{-6}
$\beta_{2,3}=\beta_{3,2}$	subadult–adult	1.75×10^{-6}

Using the model with strong heterogeneous mixing and density dependence, chosen by the model selection tests, we estimated point values for each of the elements in the WAIFW matrix according to the maximum-likelihood estimates (table 3). Among the juvenile stages, transmission with subadults comprised the greatest component of disease incidence ($\beta_{1,2}=8.99 \times 10^{-5}$), closely followed by transmission within the juvenile stage ($\beta_{1,1}=3.26 \times 10^{-5}$) and transmission between juveniles and adults ($\beta_{1,3}=2.92 \times 10^{-6}$; table 3). Subadults showed the greatest interaction with juveniles ($\beta_{2,1}=8.99 \times 10^{-5}$), followed by intra-stage interactions ($\beta_{2,2}=1.93 \times 10^{-5}$) and interactions with adults ($\beta_{2,3}=1.75 \times 10^{-6}$; table 3). Lastly, adult transmission was the greatest within the adult stage class ($\beta_{3,3}=6.42 \times 10^{-6}$) and decreased with both juveniles ($\beta_{3,1}=2.92 \times 10^{-6}$) and subadults ($\beta_{3,2}=1.75 \times 10^{-6}$; table 3). Model simulations were run with the parameter values for both density-dependent transmission (figure 2a) and frequency-dependent transmission (figure 2b).

The basic reproductive ratio, as given by the dominant eigenvalue of the next-generation matrix (equation (2.14)) using transmission estimates from the unconstrained simultaneous optimization of all parameters and

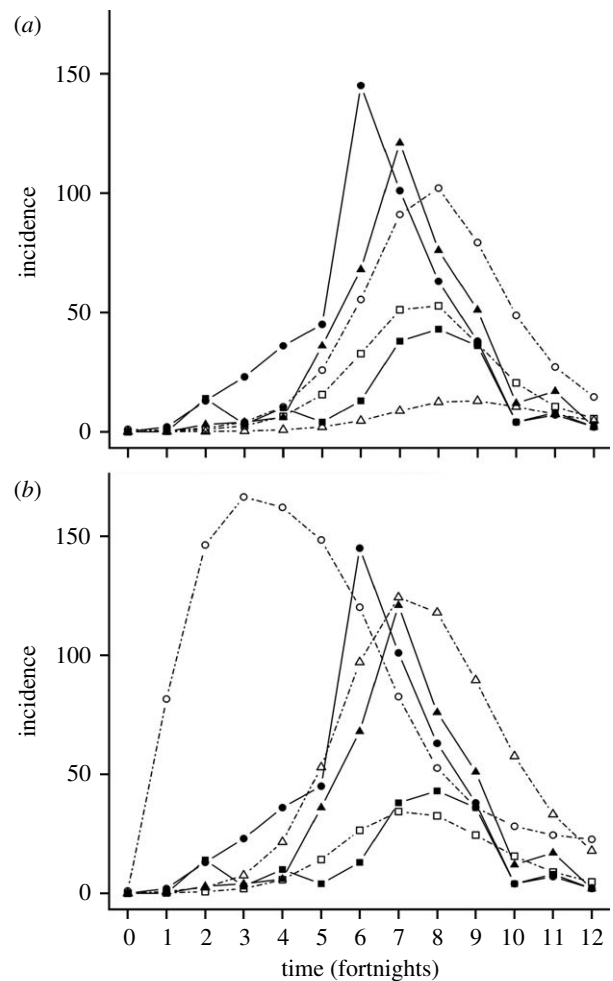


Figure 2. Strongly heterogeneous mixing model predictions for density-dependent and frequency-dependent transmissions. The best-fit model predictions for incidence, or the number of new PDV cases, are plotted with the stranded harbour seal data for the model with (a) density-dependent transmission and (b) frequency-dependent transmission (filled squares, juvenile data; filled circles, subadult data; filled triangles, adult data; open squares, juvenile model; open circles, subadult model; open triangles, adult model). Initial model conditions are plotted at time 0.

density-dependent transmission dynamics, gave an estimate of $R_0=2.03$. For the model with strong heterogeneous mixing and frequency-dependent transmission dynamics, $R_0=2.20$. Both estimates fall directly

in the range of other R_0 estimates for PDV (Swinton *et al.* 1998), particularly in the range for those in the Dutch Wadden Sea (Klepac 2007).

4. DISCUSSION

Heterogeneities—specifically, age- or stage-structured behaviour—dictate transmission of many diseases (Anderson & May 1982, 1984, 1985; Schenzle 1984), and current (as well as previous) evidence points to important stage structure in the PDV transmission among harbour seals. Harbour seals are known to behaviourally discriminate their interactions based on stage class (Wilson 1974; Sullivan 1982; Renouf & Lawson 1986, 1987; Godsell 1988; Thompson *et al.* 1989; Traut 1999). Moreover, PDV incidence has previously been shown to have signatures of stage dependence in the Dutch 2002 outbreak (Rijks *et al.* 2005). In this paper, we develop a theoretical framework to incorporate stage-structured heterogeneities in a PDV SIR model. By creating and ranking a suite of nested models ranging from a complete lack of stage structure with homogeneous mixing through various levels of heterogeneous mixing, we showed that the model with added stage structure provided a better fit for the incidence data. Overall, the model with strong heterogeneities and density-dependent (rather than frequency-dependent) transmission provided the best fit.

Using the full stage-structure model with the symmetrical β matrix, we were able to determine the elements of the WAIFW matrix from incidence data alone (table 3), illuminating both mechanisms of epidemic spread and harbour seal contact structure. We were also able to use the parameter values to recreate the epidemic in the harbour seal population (figure 2) and verify our estimates by calculating R_0 using the next-generation formalism, which resulted in a value that is acceptable for PDV ($R_0 = 2.03$ for density-dependent model, $R_0 = 2.20$ for frequency-dependent model).

Although we calculated point estimates for elements in the WAIFW matrix, these elements are highly correlated, so it is difficult to unambiguously estimate certain pairs of parameters. Perhaps this is due to model assumptions: errors in exact age determination and assuming bidirectionality in disease transmission may have confounded stage relationships. Although our method for estimating the WAIFW matrix from incidence data provides a useful tool to study heterogeneities in transmission, the method cannot capture all details. The resulting ambiguities in the parameter estimates highlight the need for additional behavioural data, without which questions on detailed mechanisms of transmission cannot be answered. For example, haul-out data (Härkönen *et al.* 1999) show that there is a temporal segregation of seals in different classes on their haul-out locations. This suggests that there might be a seasonal forcing to values in the transmission matrix, which our model ignores. To make the theoretical results relevant, information about the number and type of intra-stage and inter-stage interactions must be determined. Additional ambiguity in the models may result from the lack of a clearly distinguishable latent period, during which the infected individual is not infectious, leading to model incidence peaks that are shifted to the left. Nevertheless, these points do not

undermine the result that transmission heterogeneities play a clear role, given the LRTs (table 2). Thus, LRTs can distinguish heterogeneous mixing from homogeneous mixing in this stage-structured data.

The importance of stage-structured heterogeneity in transmission has implications for future PDV epidemics in the North Sea and for the harbour seal population structure. First, temporary or long-term change in abundance, behaviour, susceptibility and many other characteristics for any stage class may alter transmission dynamics, and therefore disease progression, through the harbour seal host populations. Our results are also important for understanding the dynamics and the mechanisms of spread of other airborne pathogens that have a mode of transmission similar to PDV. In outbreaks of other pathogens, results on heterogeneities in transmission can have implications for control strategies. For example, understanding the role that different population segments play in transmission can identify a group of individuals with the greatest transmission potential, and targeting that high-risk subgroup may be the most efficient control strategy (Anderson & May 1991). Second, transmission heterogeneities may lead to differential mortality, which in turn may affect the demographic structure of the harbour seal population in the Dutch Wadden Sea. Changes in abundance of any of the three classes may affect mating, reproduction, survival and other vital population parameters, at least in the near future.

The model with strong heterogeneous transmission and density dependence best fit the data. But in reality, frequency- and density-dependent transmissions are two endpoints in a continuum of possible transmission scenarios (McCallum *et al.* 2001). It is likely that our case study population of PDV in harbour seals in the Dutch Wadden Sea in 2002 represents a situation where the heterogeneous host population exhibits differential behaviour in terms of transmission dynamics. For example, both the juvenile and adult seals showed incidence that could be explained by the frequency-dependent model (figure 2a); this correlates to the behaviour of these seals, which are known to be relatively isolated from the rest of the herd, particularly during the pupping season. However, subadults are the most social group and appear to have incidence that is best explained by the density-dependent model (figure 2b). This suggests that, in reality, populations are heterogeneous in terms of their transmission dynamics: even within the same host population, subgroups can display differences in social behaviour and transmission dynamics.

In conclusion, stage structure clearly plays an important role in the dynamics of this epidemic. Elements in the WAIFW matrix, which provide information about transmission dynamics within and between stage classes of harbour seals, can be estimated based on stage incidence data alone. However, identifiability and uncertainty issues still exist within these parameter estimates, highlighting the need for additional behavioural data to restrict the ranges of the theoretical parameter estimations to biologically plausible and realistic values. Combining our statistical methodology with the next-generation formalism further allows us to estimate R_0 . Further analysis of epidemics among social hosts should consider abundance, behaviour, temporal

patterns in transmission and ecology of age or stage classes for more realistic models.

The authors would like to thank Bryan Grenfell for his assistance with both theory and analysis, Michael Neubert and Hal Caswell for their help with the model formulation, and Angela Luis for helpful discussions. In addition, the authors wish to thank the volunteers who covered the Dutch coast daily in search of stranded seals; the people working at the Seal Research and Rehabilitation Center in Pieterburen for logistical support; VOP Containers for providing the location to necropsy the seals; the staff of the Dutch Ministry of Agriculture, Nature and Food Quality (LNV-Noord) for providing access to and help with the centralized seal registration data; and the Common Wadden Sea Secretariat and the Trilateral Seal Expert Group for international coordination of the outbreak. L.W.P. was supported by the National Science Foundation, under the NSF Graduate Teaching Fellowship in K-12 Education (DGE-0338240). P.K. acknowledges the support of a UNESCO-I'Oréal Fellowship.

REFERENCES

- Abt, K. F. 2002 Phanologie und Populationsdynamik des Seehundes (*Phoca vitulina*) im Wattenmeer: Grundlängen zur Messung von Statusparametern vol. PhD thesis, Christian-Albrechts-Universität zu Kiel, Kiel, Germany.
- Allen, L. & van den Driessche, P. 2008 The basic reproduction number in some discrete-time epidemic models. *J. Differ. Eqns. Appl.* **14**, 1127–1147. (doi:10.1080/10236190802332308)
- Altizer, S. et al. 2003 Social organization and parasite risk in mammals: integrating theory and empirical studies. *Annu. Rev. Ecol. Evol. Syst.* **34**, 517–547. (doi:10.1146/annurev.ecolsys.34.030102.151725)
- Anderson, R. M. & May, R. M. 1982 The control of communicable diseases by age-specific immunization schedules. *Lancet* **319**, 160. (doi:10.1016/S0140-6736(82)90396-8)
- Anderson, R. M. & May, R. M. 1984 Spatial, temporal, and genetic heterogeneity in host populations and the design of immunization programs. *Math. Med. Biol.* **1**, 233–266. (doi:10.1093/imammb/1.3.233)
- Anderson, R. M. & May, R. M. 1985 Age-related changes in the rate of disease transmission: implications for the design of vaccine programmes. *J. Hyg. (Cambridge)* **94**, 365–436.
- Anderson, R. M. & May, R. M. 1991 *Infectious diseases of humans: dynamics and control*. Oxford, UK: Oxford University Press.
- Baker, J. R. 1992 The pathology of phocine distemper. *Sci. Total Environ.* **115**, 1–7. (doi:10.1016/0048-9697(92)90027-P)
- Belisle, C. J. P. 1992 Convergence theorems for a class of simulated annealing algorithms. *J. Appl. Probab.* **29**, 855–895. (doi:10.2307/3214721)
- Cosby, S. L. et al. 1988 Characterization of a seal morbillivirus. *Nature* **336**, 115–116. (doi:10.1038/336115a0)
- de Jong, M., Diekmann, O. & Heesterbeek, J. 1994 The computations of R_0 for discrete-time epidemic models with dynamic heterogeneity. *Math. Biosci.* **119**, 97–114. (doi:10.1016/0025-5564(94)90006-X)
- De Koeijer, A., Diekmann, O. & Reijnders, P. 1998 Modelling the spread of phocine distemper virus among harbour seals. *Bull. Math. Biol.* **60**, 585–596. (doi:10.1006/bulm.1997.0030)
- Diekmann, O. & Heesterbeek, J. A. P. 2000 *Mathematical epidemiology of infectious diseases: model building, analysis, and interpretation*. New York, NY: Wiley.
- Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. 1990 On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382. (doi:10.1007/BF00178324)
- Dietz, R., Heide-Jorgensen, M.-P. & Härkönen, T. 1989 Mass deaths of harbor seals (*Phoca Vitulina*) in Europe. *Ambio* **18**, 258–264.
- Dobson, A. 2004 Population dynamics of pathogens with multiple host species. *Am. Nat.* **164**, S64–S78. (doi:10.1086/424681)
- Edmunds, W. J., O'Callaghan, C. J. & Nokes, D. J. 1997 Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. *Proc. R. Soc. Lond. B* **264**, 949–957. (doi:10.1098/rspb.1997.0131)
- Ferguson, N. M., Donnelly, C. A. & Anderson, R. M. 2001 The foot-and-mouth epidemic in Great Britain: pattern of spread and impact of interventions. *Science* **292**, 1155–1160. (doi:10.1126/science.1061020)
- Ferrari, M. J., Bansal, S., Meyers, L. A. & Bjornstad, O. N. 2006 Network frailty and the geometry of herd immunity. *Proc. R. Soc. B* **273**, 2743–2748. (doi:10.1098/rspb.2006.3636)
- Fraser, C., Riley, S., Anderson, R. M. & Ferguson, N. M. 2004 Factors that make a disease outbreak controllable. *Proc. Natl Acad. Sci. USA* **101**, 6146–6151. (doi:10.1073/pnas.0307506101)
- Godsell, J. 1988 Herd formation and haul-out behavior in harbor seals (*Phoca vitulina*). *J. Zool. Lond.* **215**, 83–98.
- Grenfell, B. T., Loneragan, M. E. & Harwood, J. 1992 Quantitative investigations of the epidemiology of phocine distemper virus (PDV) in European common seal populations. *Sci. Total Environ.* **115**, 15–29. (doi:10.1016/0048-9697(92)90029-R)
- Hall, A. J., Jepson, P. D., Goodman, S. J. & Härkönen, T. 2006 Phocine distemper virus in the North and European Seas—data and models, nature and nurture. *Biol. Conserv.* **131**, 221–229. (doi:10.1016/j.biocon.2006.04.008)
- Harder, T., Willhaus, T. H., Frey, H.-R. & Liess, B. 1990 Morbillivirus infections of seals during the 1988 epidemic in the Bay of Heligoland: III. Transmission studies of cell culture-propagated phocine distemper virus in harbor seals (*Phoca vitulina*) and a grey seal (*Halichoerus grypus*): clinical, virological and serological results. *J. Vet. Med.* **37**, 641–650.
- Härkönen, T., Harding, K. C. & Lunneryd, S. G. 1999 Age- and sex-specific behaviour in harbour seals (*Phoca vitulina*) leads to biased estimates of vital population parameters. *J. Appl. Ecol.* **36**, 825. (doi:10.1046/j.1365-2664.1999.00434.x)
- Härkönen, L. et al. 2006 A review of the 1988 and 2002 phocine distemper virus epidemics in European harbour seals. *Dis. Aquat. Org.* **68**, 115–130. (doi:10.3354/dao068115)
- Huang, Y. & Rohani, P. 2006 Age-structured effects and disease interference in childhood infections. *Proc. R. Soc. B* **273**, 1229–1237. (doi:10.1098/rspb.2005.3454)
- Jensen, T., van de Bildt, M., Dietz, H. H., Andersen, T. H., Hammer, A. S., Kuiken, T. & Osterhaus, A. 2002 Another phocine distemper outbreak in Europe. *Science* **297**, 209. (doi:10.1126/science.1075343)
- Kanaan, M. N. & Farrington, C. P. 2005 Matrix models for childhood infections: a Bayesian approach with applications to rubella and mumps. *Epidemiol. Infect.* **133**, 1009–1021. (doi:10.1017/S0950268805004528)
- Klepac 2007 Interacting populations: hosts and pathogens, prey and predators. PhD thesis, MIT, Cambridge, MA.

- Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. 2005 Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359. (doi:10.1038/nature04153)
- Mahy, B. W. J., Barrett, T., Evans, S., Anderson, E. C. & Bostock, C. J. 1988 Characterization of a seal morbillivirus. *Nature* **336**, 115. (doi:10.1038/336115a0)
- McCallum, H., Barlow, N. & Hone, J. 2001 How should pathogen transmission be modelled? *Trends Ecol. Evol.* **16**, 295–300. (doi:10.1016/S0169-5347(01)02144-9)
- Newman, M. E. J. 2005 Threshold effects for two pathogens spreading on a network. *Phys. Rev. Lett.* **95**, 108 701. (doi:10.1103/PhysRevLett.95.108701)
- Osterhaus, A. D. M. E. & Vedder, E. J. 1988 Identification of virus causing recent seal deaths. *Nature* **355**, 20. (doi:10.1038/335020a0)
- Osterhaus, A. D. M. E., Uytendaele, F. G., Visser, I. K., Vedder, E. J., Reijnders, P. J., Kuiper, J. & Brugge, H. N. 1989 Seal vaccination success. *Nature* **337**, 21. (doi:10.1038/337021a0)
- Perkins, S. E., Cattadori, I. M., Tagliapietra, V., Rizzoli, A. P. & Hudson, P. J. 2003 Empirical evidence for key hosts in persistence of a tick-borne disease. *Int. J. Parasitol.* **33**, 909–917. (doi:10.1016/S0020-7519(03)00128-0)
- R Development Core Team 2006 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Renouf, D. & Lawson, J. W. 1986 Play in harbor seals (*Phoca vitulina*). *J. Zool. Lond.* **208**, 73–86.
- Renouf, D. & Lawson, J. W. 1987 Quantitative aspects of harbour seals (*Phoca vitulina*) play. *J. Zool. Lond.* **212**, 267–273. (doi:10.1111/j.1469-7998.1987.tb05989.x)
- Ries, E. H., Hiby, L. R. & Reijnders, P. J. H. 1998 Maximum likelihood population size estimation of harbour seals in the Dutch Wadden Sea based on a mark recapture experiment. *J. Appl. Ecol.* **35**, 332–339. (doi:10.1046/j.1365-2664.1998.00305.x)
- Rijks, J. M., Van de Bildt, M. W. G., Jensen, T., Philippa, J. D. W., Osterhaus, A. & Kuiken, T. 2005 Phocine distemper outbreak, The Netherlands, 2002. *Emerg. Infect. Dis.* **11**, 1945–1948.
- Schenzle, D. 1984 An age-structured model of pre- and post-vaccination measles transmission. *Math. Med. Biol.* **1**, 169–191. (doi:10.1093/imammb/1.2.169)
- Sullivan, R. M. 1982 Antagonistic behavior and dominance relationships in the harbor seals. *J. Mammol.* **63**, 544–569. (doi:10.2307/1380260)
- Swinton, J. 1998 Extinction times and phase transitions for spatially structured closed epidemics. *Bull. Math. Biol.* **60**, 215–230. (doi:10.1006/bulm.1997.0014)
- Swinton, J., Harwood, J., Grenfell, B. T. & Gilligan, C. A. 1998 Persistence thresholds for phocine distemper virus infection in harbour seal *Phoca vitulina* metapopulations. *J. Anim. Ecol.* **67**, 54–68. (doi:10.1046/j.1365-2656.1998.00176.x)
- Thompson, P. M., Fedak, M. A., McConnell, B. J. & Nicholas, K. S. 1989 Seasonal and sex-related variation in the activity patterns of common seals (*Phoca vitulina*). *J. Appl. Ecol.* **26**, 521–535. (doi:10.2307/2404078)
- Traut, I. M. 1999 Spacing among harbor seals (*Phoca vitulina vitulina*) on haul-out sites in the Wadden Sea of Niedersachsen. *Z. Sugetierkunde* **64**, 51–53.
- Trilateral Seal Expert Group 2001 Common seals in the Wadden Sea in 2001. In *Wadden Sea Newsletter 2001*, p. 3. Common Wadden Sea Secretariat.
- van den Driessche, P. & Watmough, J. 2002 Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* **180**, 32–34. (doi:10.1016/S0025-5564(02)00108-6)
- Wallinga, J., Teunis, P. & Kretzschmar, M. 2006 Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am. J. Epidemiol.* **164**, 936–944. (doi:10.1093/aje/kwj317)
- Wilson, S. 1974 Juvenile play of the common seal, *Phoca vitulina vitulina*, with comparative notes on the gray seal *Halichoerus grampus*. *Behaviour* **48**, 37–60. (doi:10.1163/156853974X00246)
- Wolfram Research 2007 *MATHEMATICA*. Champaign, IL: Wolfram Research, Inc.
- Woolhouse, M. E. J. *et al.* 1997 Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proc. Natl Acad. Sci. USA* **94**, 338–342. (doi:10.1073/pnas.94.1.338)