Canadian Science Publishing

ARTICLE

A novel approach for directly incorporating disease into fish stock assessment: a case study with seroprevalence data

John T. Trochta, Maya L. Groner, Paul K. Hershberger, and Trevor A. Branch

Abstract: When estimating mortality from disease with fish population models, common disease surveillance data such as infection prevalence are not always informative, especially for fast-acting diseases that may go unobserved in infrequently sampled populations. In these cases, seroprevalence — the proportion of fish with measurable antibody levels in their blood — may be more informative. In cases of life-long immunity, seroprevalence data require less frequent sampling intervals than infection prevalence data and can reflect the cumulative exposure history of fish. We simulation tested the usefulness of seroprevalence data in an age-structured fish stock assessment model using viral hemorrhagic septicemia virus (VHSV) in Pacific herring (*Clupea pallasii*) as a case study. We developed a novel epidemiological model to simulate population dynamics and seroprevalence data and fitted to these data in an integrated catch-at-age model with equations that estimate age- and time-varying mortality from disease. We found that simulated seroprevalence data can provide accurate estimates of infection history and disease-associated mortality. Importantly, even models that misspecified nonstationary processes in background or disease-associated mortality, but included seroprevalence data, accurately estimated annual infection and population abundance.

Résumé : Les données courantes de surveillance des maladies, par exemple sur la prévalence d'infections, n'offrent pas toujours de l'information utile pour l'estimation de la mortalité par maladie à l'aide de modèles de populations de poissons, particulièrement pour des maladies à action rapide qui pourraient passer inaperçues dans des populations échantillonnées peu fréquemment. Dans ces cas, la séroprévalence, soit la proportion de poissons présentant des concentrations mesurables d'anticorps dans le sang, peut fournir de l'information utile. Dans les cas d'immunité durable, les données de séroprévalence requièrent un échantillonnage moins fréquent que les données de prévalence d'infections et peuvent refléter les antécédents d'exposition cumulatifs des poissons. Nous avons vérifié par simulation l'utilité de données de séroprévalence dans un modèle d'évaluation de stock de poissons structuré par âge en utilisant le virus de la septicémie hémorragique virale (VSHV) chez des harengs du Pacifique (*Clupea pallasii*) comme étude de cas. Nous avons développé un nouveau modèle épidémiologique pour simuler la dynamique de la population et les données de séroprévalence et avons calé ces données sur un modèle intégré de prises en fonction de l'âge avec des équations qui estiment la mortalité par la maladie selon l'âge et variable dans le temps. Nous constatons que les données de séroprévalence simulées peuvent fournir des estimations exactes des antécédents d'infection et de la mortalité associée à la maladie. Fait à noter, même les modèles qui intègrent erronément des processus non stationnaires à la mortalité de référence ou associée à la maladie, mais qui incluent des données de séroprévalence estiment avec exactitude le taux d'infection annuel et l'abondance de la population. [Traduit par la Rédaction]

Introduction

Infectious and parasitic diseases can rapidly alter the biomass, abundance, and demographics of fish populations (Ford and Haskin 1982; Gaughan et al. 2000). These alterations arise from direct or indirect host mortalities from disease that occur in addition to background mortality and substantially reduce survival rates for a single year or across multiple years. For example, an analysis of chronic infections from mycobacteriosis in Chesapeake Bay striped bass (*Morone saxatilis*) with a multi-state mark–recapture model estimated average yearly survival of moderately and severely diseased fish to be 35% and 13%, respectively, compared to 44% in healthy fish (Groner et al. 2018). It is important for fisheries management to account for mortality from disease when it is substantial; however, few stock assessments do this even when there is strong evidence for disease impacts (Hoenig et al. 2017). The stock assessment of striped bass in Chesapeake Bay has used an infection prevalence index (the observed proportion of individuals that exhibit visible disease signs or test positive for an active infection) of mycobacteriosis with tagging data to estimate mortality from disease (Vogelbein et al. 2012). However, this application is not generalizable to acute diseases (i.e., surveillance misses an outbreak), or for host species that have low recapture rates (i.e., low power).

In other applications, mortality from disease is estimated outside stock assessment models. For example, simple logistic regression models are used to analyze disease impacts from mark–recapture data of snow crab (*Chionoecetes opilio*) off Newfoundland with bitter

Received 21 April 2021. Accepted 15 September 2021.

J.T. Trochta* and T.A. Branch. School of Aquatic and Fishery Sciences, Box 355020, University of Washington, Seattle, WA 98195, USA.
M.L. Groner. US Geological Survey, Western Fisheries Research Center, Seattle, WA 98115, USA; Prince William Sound Science Center, Cordova, AK 99574, USA.
P.K. Hershberger. US Geological Survey, Western Fisheries Research Center, Marrowstone Marine Field Station, Nordland, WA 98358, USA.
Corresponding author: John Trochta (email: john.tyler.trochta@hi.no).
*Present address: Institute of Marine Research, P.O. Box 1870, Nordnes, Bergen 5817, Norway.

crab disease and American lobster (*Homarus americanus*) in Long Island Sound with epizootic shell disease (Hoenig et al. 2017). These analyses have their drawbacks, including the need for multiple year datasets and high tag recapture rates. Additionally, they do not account for biases in selectivity. Furthermore, while it is possible for external estimates of mortality from disease to be directly input into stock assessment models, it is not straightforward to propagate uncertainty of these estimates. Ideally, disease surveillance data should be directly integrated with population data (e.g., survey abundance, age or life stage structure) and modeled with other nonstationary processes (e.g., growth and recruitment) to simultaneously estimate mortality from disease, population dynamics and population size. Therefore, alternative approaches are needed to rigorously account for and estimate mortality from disease directly within stock assessment models.

Serological assessments offer a promising epidemiological tool for disease surveillance and modeling. Serological assays test for antibodies directed against a particular agent and provide a synoptic view of prior exposure to a pathogen by measuring seroprevalence — the proportion of fish in a sample with detectable levels of circulating antibodies to the pathogen in question. Seroprevalence is particularly valuable in two common scenarios: (1) when infrequent infection prevalence surveillances do not occur concomitantly with peak infections from fast-acting pathogens (e.g., because of seasonality, and opportunistic disease surveillance) and (2) when diseased fish have higher mortality rates relative to non-diseased fish, which makes the "apparent prevalence" from surveillance an underestimate of the true infection prevalence (Heisey et al. 2006). Age-specific seroprevalence data may be used with population abundance and age structure data to estimate past mortality because long-term immunity allows for increasing seroprevalence in cohorts over time as they are exposed to infections.

The availability of seroprevalence data for diseases of wild marine fishes is a novel development. For example, assays have been recently validated for detecting antibodies for viral hemorrhagic septicemia virus (VHSV) in fish (Hart et al. 2017; Hershberger et al. 2016; Wilson et al. 2014). Age-specific serological tests, such as the indirect enzyme-linked immunosorbent assay (ELISA) and plaque neutralization test (PNT) have been successfully demonstrated for VHSV in Pacific herring (*Clupea pallasii*) and other freshwater species (Hart et al. 2017; Wilson et al. 2014). The PNT has been recently optimized to provide more sensitive and repeatable results in Pacific herring, although underestimates immunity (Hart et al. 2017). Because survivors of VHSV exposure develop long-term immunity, the exposure history of susceptible fishes should be deducible using a serological approach. By incorporating these exposure histories with population demographic data, we can deduce mortality caused by VHSV. Further, because survivors of VHSV are no longer susceptible to the disease, we can also use this approach to deduce the risk of future outbreaks (Hershberger et al. 2016).

VHSV provides an ideal case study for our analysis because its epizootiology is relatively well studied (Hershberger et al. 2016). VHSV infects various fish species in the northeast Pacific Ocean (Meyers and Winton 1995; Skall et al. 2005), where it causes periodic epizootics and fish kills (Garver et al. 2013). Mortality from VHSV and other pathogens is a leading hypothesis for limiting the recovery of Pacific herring in Prince William Sound, Alaska (Marty et al. 2003; Marty et al. 2010). This hypothesis motivated the incorporation of VHSV infection prevalence into the herring stock assessment (Marty et al. 2010; Muradian et al. 2017). Infection prevalence indices from annual fish health surveys are directly input into the Prince William Sound herring stock assessment as a linear fixed effect on natural mortality from 1994 (the first year prevalence samples were collected) to the present (Muradian et al. 2017). However, VHSV infection prevalence can easily miss short epizootics occurring outside the timing of the annual survey period and does not quantify VHS-associated mortality or survivors of the disease. In this context, VHSV infection prevalence data have limited utility for estimating annual mortality rates in wild populations (Elston and Meyers 2009), and a new disease model is needed for this herring stock assessment.

Here, we create the first fish stock assessment model that includes age-specific seroprevalence data for a herring-like fish susceptible to a fast-acting disease, similar to VHS, that confers life-long immunity via antibody production. We (1) develop an epidemiological model with annual population dynamics to simulate outbreaks of a fast-acting viral disease, (2) develop an age-structured stock assessment model that incorporates seroprevalence data to directly estimate infection rates and disease-associated mortality, and (3) evaluate the accuracy of stock assessment model estimates using simulated seroprevalence data from the epidemiological model.

Methods

We conducted a simulation analysis of a fish population model with seroprevalence data. The simulation framework consisted of an operating model with stochastic recruitment that was used to simulate disease outbreaks in a population, and an estimation model that incorporated data generated by the operating model and estimated key population parameters (Box 1; Fig. 1). To simulate a fast-acting disease, we structured both the operating and estimation models following the principles of VHSV epidemiology that have been determined from controlled experiments (Hershberger et al. 2016). The operating model has a daily time step to simulate infections of fish of different ages and disease stages; and an annual time step to simulate fish population dynamics. Key outputs of this operating model include recruitment, mortality from disease, other natural mortality, infection prevalence, and seroprevalence. Outputs from the operating model were sampled to mimic data collection during annual scientific surveys of fish populations, and included relative fish abundance, age composition, and seroprevalence by age. Simulated survey data were analyzed using an estimation model, which is loosely based on equations and estimated parameters from the age-structured stock assessment model of the Pacific herring population in Prince William Sound, Alaska (Muradian et al. 2017). The estimation model keeps track of annual age-structure, but ignores daily disease transmission; instead, annual disease infection and recovery were estimated directly. Different parameterizations of the operating and estimation models were explored to test whether disease-related parameters could be estimated under a variety of scenarios. Test scenarios included the effects of the quality of data collection through surveys, lagged transmission due to age-specific mixing of susceptible fish, and ignoring time-varying background natural mortality rates. Definitions for key modeling and epidemiological terms used in this study are provided in Box 1. Performance of the estimation models was evaluated by comparing estimates of annual spawning biomass, recruitment, and seroprevalence to the "true" values output by the operating models.

Operating model overview

We adapted a modeling framework similar to Briggs et al. (2005), in which within-year disease dynamics were modeled distinctly from between-year population dynamics. This model ran a sequence of two events within each year: (1) a fixed period of transmission, recovery, and mortality transitions occurring on daily time steps for a portion of the year and (2) a single time step that projects the total population dynamics to the start of the next year (Fig. 1). This sequence reflects a life history in which there is a fixed period conducive to disease transmission, for example during annual herring spawning aggregation events, or when seasonal environmental conditions conducive to outbreaks (e.g., temperature and salinity) are occurring. Because this is the first attempt to account for disease with seroprevalence data in a fish population dynamics model, for simplicity we did not

Box 1. Terms and definitions.

Term	Definition
Operating model	In this study, the mathematical model with equations describing the underlying epidemiology and biology of a population. Output from this model represents the true values of the population abundance and demographic rates.
Estimation model	Mathematical and statistical model with equations describing annual population dynamics, predicting data, and fitting data predictions to observations. This takes output from the operating model as input. Parameters representing annual recruitment, annual infection, recovery from disease, survey selectivity, and age-specific mixing are estimated.
Simulation analysis	Statistical analysis involving computing the errors in estimates from the estimation model with their corresponding true values from the operating model.
Replicates	Different realizations of outputs from the operating model. Each replicate has a different set of values for recruitment deviations (ε_y) and observation error (δ_y).
Recruitment deviation–deviation	The random component of the total recruitment added to the expected recruitment (in this study, this is the mean recruitment).
Observation error	The random component of the data added to each simulated observation.
Spawning biomass	The total mass (in metric tonnage) of fish in the population that are old enough to spawn.
Recruitment	The abundance of new individuals of a single age group (age 0 in this analysis) joining the population in a given year. In fished populations, recruitment generally refers to the new individuals joining the fishable population.
Infection prevalence	The proportion of a sample or population that tests positive for infection at specific point in time.
Seroprevalence	The proportion of a sample or population that tests positive for antibodies at a specific point in time.
Total population	All individuals of all ages and disease stages.
Reservoir population	The portion of the total population within which a pathogen replicates and infection is transmitted between individuals.
Total susceptible	The number of all individuals in the total population that are susceptible to infection.
Reservoir susceptible population	The number of all individuals in the reservoir population that are susceptible to infection.
Natural mortality	The average instantaneous rate that represents the removal of individuals from a population by all natural causes except disease in this study.
Survey selectivity	The probability that fish are captured by survey gear as a function of age.
Maturity	The probability of fish being a part of the spawning population as a function of age.
Age-specific mixing of susceptible	Age-specific probabilities of the total susceptible being reservoir susceptible. In this study, the probabilities are a logistic function of age.

incorporate fishing into this framework, although it is straightforward to adjust the equations to include catches.

Operating model: the fixed period of daily disease, recovery, and mortality transitions

For the period of disease transmission, we adapted a discrete-time model commonly used in epidemiological studies (e.g., Klepac and Caswell 2011; Metcalf et al. 2012; Winter et al. 2018) to project daily numbers of fish in each of eight age classes (age 0 through 7+) and three disease stages: susceptible ($S_{d,a}$), infected ($I_{d,a}$), and carrier ($C_{d,a}$), where subscript d is day and a is fish age in years. The 7+ age class (n_a) groups fish of all ages over 7 and is referred to as the plus group. This "SIC" model is similar to the classical susceptible–infected–recovered (SIR) model (Anderson and May 1979) except that the "recovered" stage is replaced by "carrier" to account for surviving individuals continuing to host low-level infections. For example, Pacific herring surviving VHSV continue to host the virus and propagate virions at low levels (Hershberger et al. 2010). Thus, both infected and carrier individuals can contribute to transmission, albeit at different rates.

The above model of transmission assumes all fish are completely mixed in a homogeneous population and have similar exposure. This is overly simplistic because populations mix heterogeneously, often by age. For example, spawning aggregations of herring school separately from pre-recruitment juveniles in isolated embayments. If we assume all newly born fish are susceptible (i.e., no maternal passive immunity), then the total population can be further categorized into two groups: a reservoir group (the population within which the pathogen replicates and infection is transmitted; Haydon et al. 2002) and a non-reservoir group. The non-reservoir group has not yet experienced pathogen exposure and, consequently, are all susceptible. We therefore allow for the modeling of the proportion of susceptible fish that mix with the reservoir population as follows:

$$S_{y,d=1,a} = v_a T_{y,d=1,a} \quad (1)$$

where $S_{y,d=1,a}$ is the susceptible numbers-at-age in the reservoir group (hereinafter reservoir susceptibles) on day 1, v_a is the fraction of the total susceptible numbers at age a mixing with the reservoir population, and $T_{y,d=1,a}$ is the total susceptible numbers-at-age on day 1 (hereinafter total susceptibles). The quantity v_a is specified by a logistic curve that accounts for when young fish mix with the reservoir population and first become exposed to natural transmission:

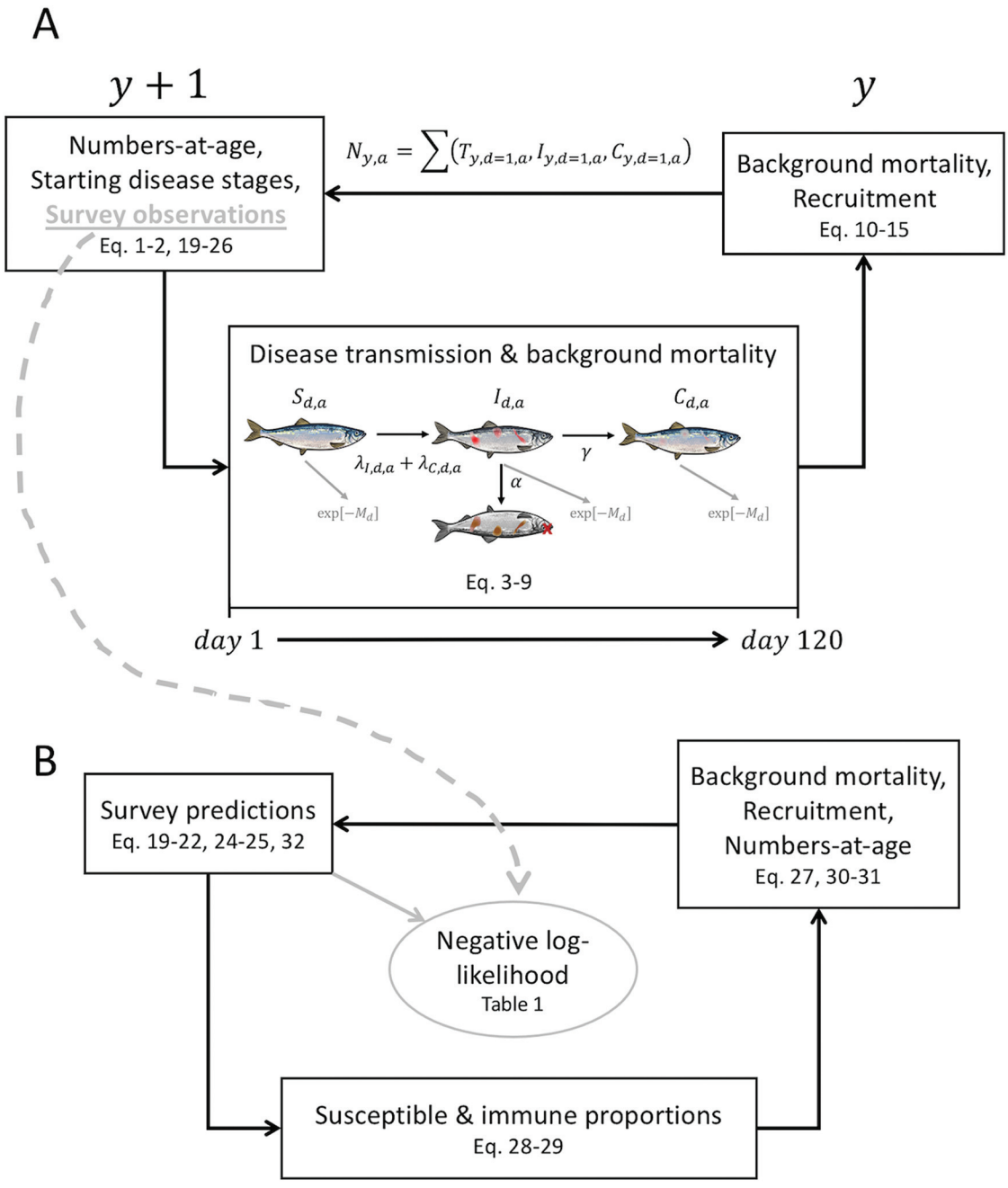
$$v_a = \frac{1}{1+\exp\left[\frac{-\ln(19)\left(a-a_{50}^{\text{mix}}\right)}{\left(a_{95}^{\text{mix}}-a_{50}^{\text{mix}}\right)}\right]} \quad (2)$$

where a_{50}^{mix} and a_{95}^{mix} are the ages at which 50% and 95% mix with the reservoir susceptibles, respectively. For our analysis, we set $a_{50}^{\text{mix}} = 3$ and $a_{95}^{\text{mix}} = 4$ to match survey selectivity so that disease outbreaks primarily occur in the observed population (i.e., a scenario with optimal disease information and parameters that are similar to those from Prince William Sound herring; Muradian et al. 2017). We also evaluate estimation performance when age-specific mixing is ignored and when mixing occurs at earlier ages ($a_{50}^{\text{mix}} = 1$; see "Simulation scenarios").

The reservoir numbers-at-age-and-stage on day d were coded using a single vector for each possible age and stage combination, $\boldsymbol{n}(d)$:

$$\boldsymbol{n}(d) = (S_{d,0}, I_{d,0}, C_{d,0}, S_{d,1}, I_{d,1}, C_{d,1}, \ldots, C_{d,n_a}) \quad (3)$$

Fig. 1. Flow diagram of the operating model and estimation model used in simulations. Panel A shows the sequence of equations in the operating model that starts with numbers of fish in each age and disease stage as projected from the previous year (y). These age- and stage-specific numbers ($S_{d,a}$, $I_{d,a}$, $C_{d,a}$) are projected on daily time steps using transition probabilities ($\lambda_{I,d,a}$, $\lambda_{C,d,a}$, α, γ) and background mortality ($\exp(-M_d)$) within a discrete-time disease transmission model (terms are defined in Table A1). After 120 days of transmission, numbers are projected to the next year ($y + 1$) after accounting for constant background mortality and recruitment in the intervening time period. Panel B shows a similar sequence of equations in the estimation model. The key difference between the two models is the bottom box; the estimation model calculates annual susceptible and immune numbers from estimated annual infection and recovery parameters. Parameters are estimated by fitting survey predictions from the estimation model to survey observations simulated from the operating model (linked by the gray dashed line), and fitting is done by minimizing the negative log-likelihood functions in Table 1. The in-text equations are referenced from the relevant processes labeled in each box.



To project reservoir stage-specific numbers of a single age group forward one day ($S_{d+1,a}$, $I_{d+1,a}$, $C_{d+1,a}$), the current day's stage specific number vector was multiplied by an age and day-specific disease transition matrix ($\mathbf{T}_{d,a}$) and by the decay due to the background natural mortality (M):

$$(S_{d+1,a}, I_{d+1,a}, C_{d+1,a})^T = (S_{d,a}, I_{d,a}, C_{d,a})^T \cdot \mathbf{T}_{d,a} \cdot \exp(-M_d) \quad (4)$$

Background natural mortality is assumed constant ($M_d = 6.85 \times 10^{-4}$ day^{-1}, which equates to 0.25 year^{-1}) for all ages, and $(S_{d,a}, I_{d,a}, C_{d,a})^T$ is a transposed vector. The non-reservoir susceptible

numbers ($1 - v_a T_{y,d=1,a}$) were not included in eqs. 3 and 4 because these numbers do not factor into transmission and were projected outside of disease transitions (eq. 10). Natural mortality can be constant in all simulated years or can be lognormal over time about an expected value M (see "Simulation scenarios"). Descriptions of all indices, variables, and parameters from the operating and estimation models are given in Table A1 of Appendix A.

The transition matrix $\mathbf{T}_{d,a}$ contains the transition probabilities from one disease stage (in columns) to another (rows) with the order of stages being reservoir susceptible, infected, and carrier:

$$\mathbf{T}_{d,a} = \begin{bmatrix} 1-\lambda_{I,d,a}-\lambda_{C,d,a} & 0 & 0 \\ \lambda_{I,d,a}+\lambda_{C,d,a} & 1-\alpha-\gamma & 0 \\ 0 & \gamma & 1 \end{bmatrix} \quad (5)$$

where $\lambda_{I,d,a}$ is the probability of infection from actively infected ($S_{d,a}$) fish; $\lambda_{C,d,a}$ is the probability of infection from carrier ($C_{d,a}$) fish; γ is the probability of recovering from infection; and α is the probability of mortality from disease. Dead fish were ignored in transitions because they do not factor into any other equations of this model. The recovery (0.071 day^{-1}) and mortality from disease (0.030 day^{-1}) rates used in the simulations were calculated from a laboratory infection experiment (Table A1; Hershberger et al. 2013). In these trials, age-0 specific pathogen free (SPF) Pacific herring were exposed to waterborne VHSV (approximately 2000 pfu·mL^{-1}) for 1 hour. The experiment was conducted at 11 °C. For most simulation scenarios, we assumed that the rates for recovery and mortality from disease are independent of age, day, and year, but it is straightforward to adapt the model to allow mortality to vary by day and year. Annually varying rates bounded by values under cold (8 °C) and hot (15 °C) water conditions from these same experiments (recovery = 0.055–0.192 day^{-1} and mortality from disease = 0.025–0.058 day^{-1}) were explored as sensitivities amongst our simulation scenarios.

The probabilities of infection ($\lambda_{I,d,a}$ and $\lambda_{C,d,a}$) were computed using the function below:

$$\lambda_{I,d,a} = 1 - \exp\left[-\sum_j \beta_I I_{d,j} \frac{S_{d,j}}{\sum \boldsymbol{n}(d)}\right] \quad (6)$$

$$\lambda_{C,d,a} = 1 - \exp\left[-\sum_j \beta_C C_{d,j} \frac{S_{d,j}}{\sum \boldsymbol{n}(d)}\right] \quad (7)$$

where β_I and β_C are the mean daily per capita rates of viral transmission from one infected or one carrier individual, respectively, at age j to all encountered reservoir susceptible fish at age a. The form of eqs. 6 and 7 assumes frequency dependent transmission (i.e., that after exceeding a population threshold transmission does not increase with total population size). While further investigation into the transmission function is warranted for this and most marine diseases (Murray 2009), frequency dependent transmission is reasonable for fish like herring that school at high densities regardless of their abundance (sensu MacCall 1990). Therefore, the number and proximity of individual interactions likely remain similar as population sizes change (Murray 2009).

Only β_I and β_C need to be specified to fully define $\lambda_{I,d,a}$ and $\lambda_{C,d,a}$. For simplicity, β_I and β_C were assumed constant across ages and years, but with $\beta_C \ll \beta_I$. Reasonable values of β_I and β_C were found using projections of the operating model that resulted in irregular outbreaks across years, but were potentially rapid (on the order of weeks) and large (infecting more than 25% of the population) when they occurred. This emulates VHSV outbreaks in herring populations (Hershberger et al. 2016). In addition, the values chosen had to allow for transmission to be maintained covertly (i.e., very low infection prevalence rates in some years) by carrier fish. These conditions were met with $\beta_I = 0.01$ day^{-1} and $\beta_C = 0.000001$ day^{-1}.

For the operating model, we chose the length of the fixed period for disease transitions to be 120 days, similar to a typical four-month period of mass schooling in herring populations (Hay et al. 2001). This fixed period is long enough for rapid epizootics to occur, much like VHSV epizootics that start and subside within a matter of weeks (Hershberger et al. 2016). Inclusion of longer fixed periods (up to 365 days) had negligible impacts on the simulated disease and population dynamics.

At the end of the fixed period ($d_{\text{end}} = 120$), all infected individuals were assumed to recover or die (Fig. 2B) without causing any further transmission in the intervening time between the current year and the next year, resulting in this transition matrix:

$$\mathbf{T}_{d_{\text{end}},a} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \gamma/(\alpha+\gamma) & 1 \end{bmatrix} \quad (8)$$

Equation 8 was used to project stage-specific numbers one more day:

$$\left(S_{y,d_{\text{end}}+1,a}, I_{y,d_{\text{end}}+1,a}, C_{y,d_{\text{end}}+1,a}\right)^T = \left(S_{d_{\text{end}},a}, I_{d_{\text{end}},a}, C_{d_{\text{end}},a}\right)^T \cdot \mathbf{T}_{d_{\text{end}},a} \quad (9)$$

which was used to calculate the starting stage structure of the next period of transitions.

Operating model: the single time step projecting numbers to the next year

The total susceptible fish at the start of the next year ($T_{y+1,d=1,a}$) included all new fish to the population, which was modeled as a stationary mean with serially correlated deviation. Thus recruitment was modeled as independent of parental abundance, because stock–recruitment relationships are seldom found in forage fish (Szuwalski et al. 2015, 2019). There is no evidence to suggest recruitment is affected by VHSV through reduced reproductive success of infected spawners. The remaining susceptible fish from the current year adjusted for removals from natural mortality that occurred since the 120th day of the previous year:

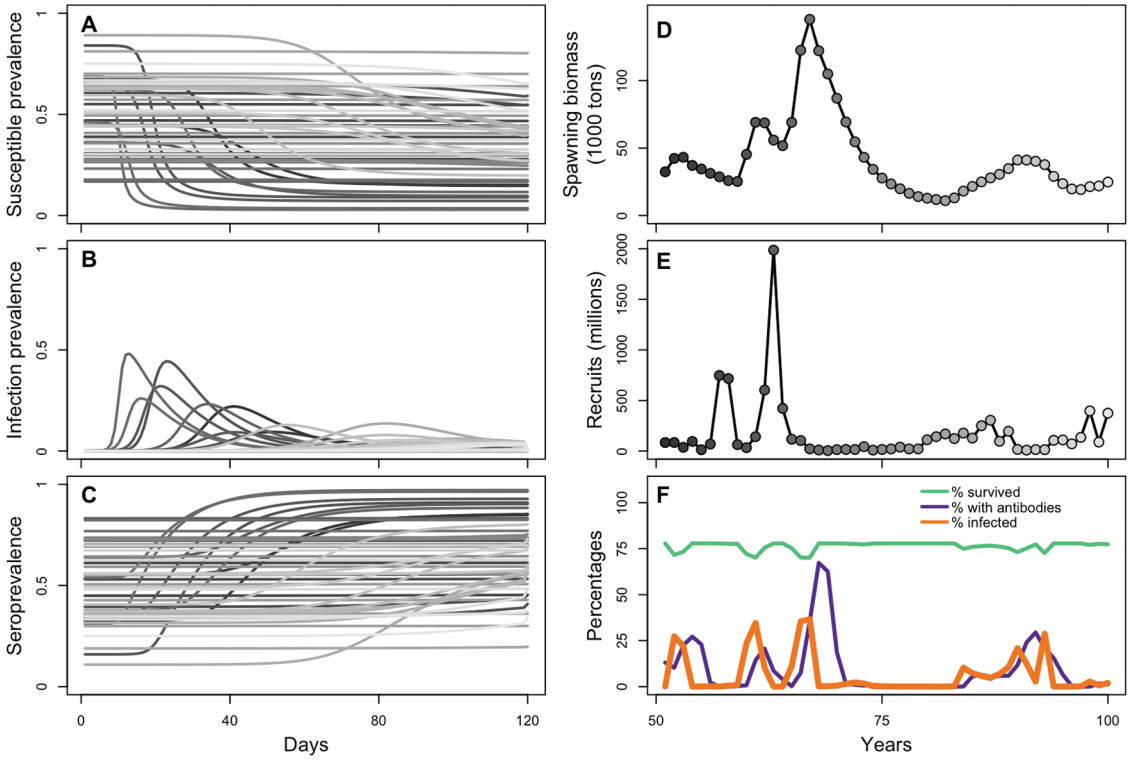
$$T_{y+1,d=1,a+1} = \begin{cases} \overline{R}\exp\left(\rho\varepsilon_{R,y} + \sqrt{1-\rho^2}\varepsilon_{R,y+1} - 0.5\sigma_R^2\right) & \text{if} \quad a = 0 \\ (1-v_a)T_{y,d=1,a}\exp(-M_d \cdot 365) + S_{y,d_{\text{end}}+1,a}\exp(-M_d \cdot \Delta t) & \text{if} \quad 0 < a < n_a \end{cases} \quad (10)$$

where $\overline{R}$ is average recruitment in numbers of age 0 fish (millions); ρ is autocorrelation in annual recruitment; $\varepsilon_{R,y}$ are the lognormal annual deviations in recruitment in year y; σ_R is the standard deviation of the annual recruitment deviations; and Δt is equal to 365 – d_{end}, the length of time between the end of the current season and the start of the next.

The reservoir susceptible numbers in the next year ($S_{y+1,d=1,a+1}$) were calculated from eq. 1. The number of infected fish was zero in the next year:

$$I_{y+1,d=1,a+1} = 0 \quad (11)$$

Fig. 2. Population and epidemiological dynamics over 50 years from one replicate of the operating model. Daily disease stage transitions are shown as the proportions of the overall population that are (A) susceptible, (B) infected, or (C) recovered and now carriers on each day of the fixed transmission period. Each line is a different year. Gray shading denotes the corresponding years between plots A–C and D–E. The annual population level dynamics are shown in the right column including (D) annual spawning biomass, (E) stochastic recruits as the number of age-0 fish, and (F) the realized population survival (the annual percentage that survive), seroprevalence (the annual percentage that are currently immune), and infection incidence (the percentage that becomes infected) of the population (excluding the plus age group) in each year. [Colour online.]



There were zero carriers among newly spawned age-0 fish and among older fish the number of carriers ($C_{y+1,d=1,a+1}$) was the number that survived natural mortality from the current year:

$$C_{y+1,d=1,a+1} = \begin{cases} 0 & \text{if} \quad a = 0 \\ C_{y,d_{\text{end}}+1,a}\exp(-M_d \cdot \Delta t) & \text{if} \quad 0 < a < n_a \end{cases} \tag{12}$$

Stage-specific numbers in the plus group next year were adjusted for the remaining numbers of the plus-age group from the current year (Table A2):

$$T_{y+1,d=1,n_a} = \exp(-M_d \cdot 365)\left[(1-\nu_{n_a-1})T_{y,d=1,n_a-1} + (1-\nu_{n_a})T_{y,d=1,n_a}\right] + \exp(-M_d \cdot \Delta t)\left(S_{y,d_{\text{end}}+1,n_a-1} + S_{y,d_{\text{end}}+1,n_a}\right) \tag{13}$$

$$I_{y+1,d=1,n_a} = 0 \tag{14}$$

$$C_{y+1,d=1,n_a} = \exp(-M_d \cdot \Delta t)\left(C_{y,d_{\text{end}}+1,n_a-1} + C_{y,d_{\text{end}}+1,n_a}\right) \tag{15}$$

The total numbers-at-age ($N_{y,a}$) is the sum of stage-specific numbers (eqs. 10–15) in each age group, or $\sum\left(T_{y,d=1,a}, I_{y,d=1,a}, C_{y,d=1,a}\right)$.

The age structure in the first year of the operating model was initialized at equilibrium:

$$N_{y=1,a=0} = \bar{R} \tag{16}$$

$$N_{y=1,a+1} = N_{y=1,a}\exp(-M_d \cdot 365) \qquad \text{if} \quad 0 < a < n_a \tag{17}$$

$$N_{y=1,n_a} = N_{y=1,n_a-1}\left[\frac{\exp(-M_d \cdot 365)}{1-\exp(-M_d \cdot 365)}\right] \tag{18}$$

where $N_{y=1,a}$ is in millions of fish. In the first year, each age group was assumed to start with 90% total susceptible, 0% infected, and 10% carriers. The operating model was run with annual recruitment as the only stochastic component (and natural mortality in one scenario — see "Simulation scenarios"), for 50 years to cycle out initial value effects in every individual simulation, and then for an additional 50 years for data generation and parameter estimation. We used R (R Core Team 2020) to code the operating model and conduct the analyses presented herein. All code and results are available in a public repository (https://github.com/johnt23/antibody_sim).

Data simulation

We generated time series of abundance, age composition, and seroprevalence data from the operating model for input into the estimation model (Fig. 1). Spawning biomass (SB_y) was calculated at the start of the model year from

$$SB_y = \sum_{a=0}^{n_a} m_a N_{y,a} w_a \tag{19}$$

where m_a is the maturity ogive and w_a is weight-at-age (constant across time). We make a simplifying assumption that m_a is equal to age-specific selectivity of the survey, s_a, because spawning aggregations of herring are typically sampled (i.e., all spawning biomass and only spawning biomass are available to surveys as

assumed for Prince William Sound herring; Trochta and Branch 2021). The curve for s_a was calculated from a logistic function:

$$(20) \qquad s_a = \frac{1}{1+\exp\left[\frac{-\ln(19)\left(a-a_{50}^{\text{survey}}\right)}{\left(a_{95}^{\text{survey}}-a_{50}^{\text{survey}}\right)}\right]}$$

where a_{50}^{survey} and a_{95}^{survey} are the ages at which 50% and 95% are selected, respectively.

The true spawning biomass (SB_y) was used to simulate the observed survey biomass index with observation error (B_y) as follows:

$$(21) \qquad B_y = q\text{SB}_y\exp\left(\delta_y - 0.5\sigma_B^2\right)$$

In eq. 21, q is the survey scalar, and δ_y is annual lognormal error in survey biomass indices where $\delta_y \sim \text{N}(0, \sigma_B)$. Survey age-compositions ($\Theta_{y,a}$) were generated from a multinomial with mean proportions obtained from

$$(22) \qquad \Theta_{y,a} = \frac{s_a N_{y,a}}{\sum_{a=0}^{n_a} s_a N_{y,a}}$$

To generate annual seroprevalence values, age compositions of immune fish ($\widetilde{\Omega}_{y,a}$) were calculated from

$$(23) \qquad \widetilde{\Omega}_{y,a} = \frac{C_{y,d=1,a}}{N_{y,a}}$$

which is the proportion of carrier fish in the total population at start of the current year (eq. 12). Age-specific seroprevalence, or the number of fish with antibodies ($A_{y,a}^{+}$) was assumed to be assayed from the same samples as the age composition survey (i.e., same survey selectivity s_a):

$$(24) \qquad A_{y,a}^{+} = s_a\widetilde{\Omega}_{y,a}N_{y,a}$$

Because fish that are not immune were included in these samples, the total number of fish failing to test positive for antibodies is

$$(25) \qquad A_{y,a}^{-} = s_a\left(1-\widetilde{\Omega}_{y,a}\right)N_{y,a}$$

To represent all mutually exclusive outcomes of the sampled seroprevalence, numbers of fish with and without antibodies in each age were combined:

$$(26) \qquad \boldsymbol{A_{y,p}} = \{A_{y,a=0}^{+}, A_{y,a=0}^{-}, A_{y,a=1}^{+}, A_{y,a=1}^{-}, \ldots, A_{y,a=n_a}^{+}, A_{y,a=n_a}^{-}\}$$

Mean proportions of each element of this vector (i.e., $\boldsymbol{A_{y,p}}/\sum_{p=1}^{2\cdot n_a}\boldsymbol{A_{y,p}}$) were then used to generate multinomial samples used as seroprevalence data. The multinomial assumption meant samples were informed by the age structure as well as the seroprevalence of each age group. Multinomial sample sizes for survey age-composition and seroprevalence data were the same in all years.

All simulated data were generated from spawning numbers-at-age immediately before transmission (i.e., on day 1 of 120), after accounting for survey selectivity.

Estimation model

The estimation model was based on a standard age-structured stock assessment model that annually projects the spawning numbers ($\hat{N}_{y,a}$) and biomass ($\hat{\text{SB}}_y$), and recruits ($\hat{N}_{y,a=0}$). It mirrored the population dynamics of the operating model, but simplified disease dynamics into annual groups ignoring daily stage transitions (Fig. 1B). Instead, the model calculated the proportion of each age class that evades infection each year $(1-v_a)E_{y,a}\omega_y$, where v_a is the age-specific mixing of susceptible fish with the reservoir group, $E_{y,a}$ is the proportion of fish that were susceptible prior to infection, and ω_y is the year-specific proportion of susceptible fish that are infected in the current year. The model estimated μ, the proportion of infected fish that recover and become carriers. A key feature to this model is that ω_y is the same for all ages in the reservoir group, which implicitly assumes all ages share the same underlying transmission function. The proportion of fish that evade or survive infection is $1-v_aE_{y,a}\omega_y(1-\mu)$ (see Appendix A for derivation), and hence:

$$(27) \qquad \hat{N}_{y+1,a+1} = \exp(-M_y)\left[1-v_aE_{y,a}\omega_y(1-\mu)\right]\hat{N}_{y,a} \quad 0 < a < n_a$$

where M is annual background mortality (M_y = 0.25 year^{-1} in all years). The estimable disease-related parameters are the annual proportion of susceptible fish that become infected, ω_y, and the time-invariant μ for the proportion of infected fish that recover. Both v_a and $E_{y,a}$ are derived (eqs. 2 and 29), but some values of v_a are fixed in some scenarios to improve model convergence. It was assumed that no fish remain infected between years (eq. 27), which is reasonable for VHSV in herring because outbreaks typically occur rapidly and survivors clear their infections very quickly (Hershberger et al. 2016).

Age-specific immunity ($\widetilde{\Omega}_{y+1,a+1}$) in the next year is the proportion of all fish (susceptible or recovered) that are immune because they were infected with, and recovered from, the virus in or prior to year y (see Appendix A for derivation):

$$(28) \qquad \widetilde{\Omega}_{y+1,a+1} = \begin{cases} 0 & \text{if} \quad a = 0 \\ \dfrac{\widetilde{\Omega}_{y,a} + v_aE_{y,a}\omega_y\mu}{1-v_aE_{y,a}\omega_y(1-\mu)} & \text{if} \quad 0 < a < n_a \end{cases}$$

These immune proportions were used to calculate the proportion that remain susceptible in each cohort in the current year ($E_{y,a}$):

$$(29) \qquad E_{y,a} = \begin{cases} 1 & \text{if} \quad a = 0 \\ 1-\widetilde{\Omega}_{y,a} & \text{if} \quad 0 < a < n_a \end{cases}$$

Separate equations for the plus group ($a = n_a$) adjusted for the existing numbers in n_a from year y were created and are shown in Table A2.

Additional parameters estimated by the model were the parameters of the logistic function for survey selectivity (a_{50}^{survey} and a_{95}^{survey}), lognormal annual recruitment deviations in recruitment (ε_y), mean recruitment ($\overline{R}$), and the lognormal standard deviation in recruitment (σ_R). In some scenarios, parameters of the logistic function for the proportions of young fish mixing with the reservoir population (a_{50}^{mix} and a_{95}^{mix}) were also estimated. Recruitment was modeled with the same parameterization used in the operating model (eq. 10), except that recruitment in the first year was estimated:

$$(30) \qquad \hat{N}_{y+1,a=0} = \begin{cases} \overline{R}\exp\left(\varepsilon_y - 0.5\sigma_R^2\right) & \text{if} \quad y = 1 \\ \overline{R}\exp\left(\rho\varepsilon_y + \sqrt{1-\rho^2}\varepsilon_{y+1} - 0.5\sigma_R^2\right) & \text{if} \quad y > 1 \end{cases}$$
$$\text{where} \quad \varepsilon_y \sim \text{N}(0, \sigma_R^2)$$

Numbers-at-age in the first year of observations were modeled with estimated lognormal deviations (ε_y) around an estimated mean abundance ($\overline{N}$) and σ_R:

$$(31) \qquad \hat{N}_{y=1,a} = \overline{N}\exp\left(\varepsilon_y - 0.5\sigma_R^2\right) \qquad \text{where} \quad \varepsilon_y \sim \text{N}(0, \sigma_R^2)$$

Annual recruitment and initial numbers-at-age were treated as random effects (equations in Table 1) that were integrated out in Template Model Builder (TMB) (Kristensen et al. 2016) using the

Table 1. Equations of components from the objective function in the estimation model.

Equation	Description
$\delta_N = (n_a - 1)\ln(\sigma_R) + \frac{\sum_y^{n_a-1} \varepsilon_y^2}{2\sigma_R^2}$	Penalty on numbers-at-age deviations in the first year ($y = 1$)
$\delta_R = n_y \ln(\sigma_R) + \frac{\sum_y^{n_y} \varepsilon_y^2}{2\sigma_R^2}$	Penalty on recruitment deviations
$\ell(B) = n_y \ln(\sigma_B) + \frac{\sum_y^{n_y} \left[\ln(\hat{B}_y) - \ln(B_y)\right]^2}{2\sigma_B^2}$	Negative log-likelihood of biomass survey
$\ell(\Theta) = \sum_y^{n_y} \left[Z_{\Theta,y} \sum_a^{n_a} \Theta_{y,a} \ln\left(\frac{\hat{\Theta}_{y,a}}{\Theta_{y,a}}\right)\right]$	Log-likelihood of fishery-dependent age composition
$\ell(\hat{A}) = \sum_y^{n_y} \sum_a^{n_a} \left[-k_{y,a}\ln(\hat{p}_{y,a}) - \left(N_{y,a}^{\text{sero}} - k_{y,a}\right)\ln\left(1-\hat{p}_{y,a}\right)\right]$	Log-likelihood of antibody prevalence
$\mathcal{L} = \delta_N + \delta_R + \ell(B) - \ell(\Theta) - \ell(\hat{A})$	Objective function

Note: Terms are defined in Table A1.

Laplace approximation, which allowed estimation of their respective standard deviations.

The log-likelihood functions for the biomass and age composition survey data were derived from the same distributions used to simulate these data from the operating model (e.g., $\hat{B}_y$ is the predicted survey biomass from the estimation model and counterpart to B_y, the simulated data). For the age-specific seroprevalence data, a binomial likelihood was used. We chose not to use a multinomial distribution (that was used to simulate seroprevalence) because this would result in age structure information in the seroprevalence samples (i.e., the proportions of the total sample in each age) being counted twice in model fitting — once in the proportion-at-age data and again in the seroprevalence-at-age data. In addition, using binomial likelihood allows for cases where seroprevalence sample sizes are increased during sampling to ensure larger sample sizes from rare age groups. The negative log-likelihood component for the seroprevalence data sums over the individual binomial likelihoods for each seroprevalence value for each age in each year (Table 1):

$$\ell(\hat{A}) = \sum_{y=1}^{n_y} \sum_{a=0}^{n_a} \left[-k_{y,a}\ln(\hat{p}_{y,a}) - \left(N_{y,a}^{\text{sero}} - k_{y,a}\right)\ln\left(1-\hat{p}_{y,a}\right)\right] \quad (32)$$

where $k_{y,a}$ is the observed number of fish testing positive for antibodies in age a and year y, $\hat{p}_{y,a}$ is the predicted proportion of seropositive fish in age a and year y ($\hat{p}_{y,a} = \hat{A}_{y,a}^{+} / \left(\hat{A}_{y,a}^{+} + \hat{A}_{y,a}^{-}\right)$ from eqs. 24 and 25), and $N_{y,a}^{\text{sero}}$ is the total number sampled in each age a and year y.

Predictions of the fitted data ($\hat{B}_y$, $\hat{\Theta}_{y,a}$, and $\hat{\mathbf{A}}_{\mathbf{y},\mathbf{p}}$) were calculated from the same equations used for data simulation (eqs. 20–22, 24–26). The sampling CV for the biomass index (σ_B), and effective sample sizes for both age-composition ($Z_{\Theta,y}$) and seroprevalence data ($Z_{A,y}$) were inputs to the estimation model, and were assumed to be known exactly. Spawning biomass ($\hat{\text{SB}}_y$) was estimated using $\hat{N}_{y,a}$ in eq. 19. The estimation model was implemented in TMB (Kristensen et al. 2016) and optimized with "nlminb" in R (R Core Team 2020), and is also available in the public repository (https://github.com/johnt23/antibody_sim).

Simulation scenarios

We explored the robustness of our estimation model to a variety of scenarios that varied in assumptions about fish biology, epidemiology, and sampling protocols, both within the operating and estimation models (Table 2). In each following scenario, 500 replicates of the operating model were run, where each replicate had different randomly generated values for the recruitment deviations (ε_y) and observation errors (e.g., δ_y). The random seeds were preserved amongst scenarios so that scenarios were comparable. Then the estimation model was fitted to the data generated from the operating model to determine which parameters could be estimated accurately.

Ignore disease

This simulation reflected conventional stock assessments where disease is not included. Seroprevalence data were not fit, and infection and recovery parameters were ignored in the estimation model.

Incorporate infection prevalence

This simulation matched the method currently used to estimate annual mortality from VHSV in Prince William Sound herring stock assessments: instead of using antibody data (estimating past disease), only the prevalence of VHSV infection (# VHSV positive / # Total in a single sample) was included as an annual index of additional natural mortality in the stock assessment (Muradian et al. 2017):

$$\exp\left[-\left(M_y + \kappa\overline{\Psi}_y\right)\right] \quad (33)$$

where κ is an estimated parameter that scales infection prevalence to natural mortality and $\overline{\Psi}_y$ is a year-specific infection prevalence index generated from the average of three samples collected each year. Seroprevalence data were ignored. To simulate this scenario within the operating model, the proportion of infected fish in the population available to sampling on a given day d was computed:

$$\Psi_{y,d} = \frac{\sum_{a=0}^{n_a} s_a I_{y,d,a}}{\sum_{a=0}^{n_a} s_a N_{y,d,a}} \quad (34)$$

Equation 34 was applied to numbers-at-age-and-stage on three random days within a 30-day timeframe during the 120-day transmission season, mimicking clustered sampling. True infection prevalence values were used as the mean probability in a binomial distribution to randomly generate virus-positive samples from a sample size of 60 per day.

Incorporate seroprevalence

In this simulation, the estimation model was fit to age-specific seroprevalence data with large age-specific sample sizes (200 fish sampled per year) and reflected a well-informed scenario for incorporating seroprevalence data into stock assessment. The assumed age-specific proportions mixing with the reservoir

Table 2. Specifications of the operating and estimation models for simulation scenarios.

	Operating model				Estimation model				
Name	Age comp/ antibody sample size	Mixed proportion logistic parameters	Background mortality (M_y)	Mortality and recovery probabilities (α, γ)	Infection probability (ω_y)	Recovery probability (μ)	Mixed proportion parameters	Fit antibody data	Input infection prevalence
Ignore disease	200/NA	$a_{50}^{mix} = 3$ $a_{95}^{mix} = 4$	Constant	Constant	Fix ($\omega_y = 0$)	Fix ($\mu = 0$)	NA	No	No
Incorporate infection prevalence	200/NA	$a_{50}^{mix} = 3$ $a_{95}^{mix} = 4$	Constant	Constant	Fix ($\omega_y = 0$)	Fix ($\mu = 0$)	NA	No	Yes
Incorporate seroprevalence	200/200	$a_{50}^{mix} = 3$ $a_{95}^{mix} = 4$	Constant	Constant	Est	Est	Est v_a for ages 2–4; fix $v_a = 0$ if $a < 2$, $v_a = 1$ if $a > 4$	Yes	No
Small sample size	20/20	$a_{50}^{mix} = 3$ $a_{95}^{mix} = 4$	Constant	Constant	Est	Est	Same as "Incorporate seroprevalence"	Yes	No
Early mixing of susceptible fish	200/200	$a_{50}^{mix} = 1$ $a_{95}^{mix} = 4$	Constant	Constant	Est	Est	Est v_a for ages < 5; fix $v_a = 1$ if $a > 4$	Yes	No
Age-specific mixing ignored	200/200	$a_{50}^{mix} = 3$ $a_{95}^{mix} = 4$	Constant	Constant	Est	Est	Fix $v_a = 1$	Yes	No
Time-varying background mortality	200/200	$a_{50}^{mix} = 3$ $a_{95}^{mix} = 4$	$\ln(M_y) \sim N(\ln(0.25), 0.25^2)$	Constant	Est	Est	Same as "Incorporate seroprevalence"	Yes	No
Time-varying background mortality and ignore disease	200/NA	$a_{50}^{mix} = 3$ $a_{95}^{mix} = 4$	$\ln(M_y) \sim N(\ln(0.25), 0.25^2)$	Constant	Fix ($\omega_y = 0$)	Fix ($\mu = 0$)	NA	No	No
Time-varying disease mortality and recovery	200/200	$a_{50}^{mix} = 3$ $a_{95}^{mix} = 4$	Constant	$\alpha_y \sim U(0.025, 0.058)$ $\gamma_y \sim U(0.055, 0.192)$	Est	Est	Same as "Incorporate seroprevalence"	Yes	No

Note: Gray shaded boxes highlight differences from the "Ignore disease" scenario. Terms are defined in Table A1.

population matched the maturity and survey selectivity assumed in the operating model ($v_a = s_a$).

Small sample size

This simulation was similar to "Incorporate seroprevalence" but with a small sample size (20) for seroprevalence and age composition data to reflect a data-limited scenario for stock assessment.

Early mixing of susceptible

This simulation was also similar to "Incorporate seroprevalence" except we examined estimation performance when fish had delayed exposure to infection (i.e., age-specific mixing with the reservoir population) that occurred before they were observed. In the operating model, susceptible fish mixed with the reservoir population at greater proportions starting at age 0 ($a_{50}^{mix} = 1$ year, $a_{95}^{mix} = 4$ years).

Age-specific mixing ignored

This scenario was similar to "Early mixing of susceptible", except that it explored the consequences of model misspecification in ignoring age-specific mixing with the reservoir population. Specifically, the estimation model assumed that all ages were completely mixed in the reservoir population and schooling together ($v_a = 1$).

Time-varying background natural mortality

This scenario was similar to "Incorporate seroprevalence" except that the operating model included time-varying background mortality from other natural sources that change over time, instead of constant natural mortality. Time-varying background mortality (changing M_y) was modeled as random lognormal deviations about the expected value (Table 2), but the estimation model did not estimate time-varying background mortality while incorporating seroprevalence data.

Time-varying background natural mortality and ignore disease

In this scenario, the operating model from the "Time-varying background natural mortality" scenario was paired with an estimation model that ignored disease to explore model misspecification in both background mortality and mortality from disease.

Time-varying disease mortality and recovery

This scenario was similar to "Incorporate seroprevalence" except that recovery rates and mortality from disease varied from one year to the next. Recovery (γ_y) and mortality from disease (α_y) transition probabilities were stochastically drawn from uniform distributions in the operating model for each year. VHSV infections depend on environmental conditions, so the uniform distributions were bounded by rates observed in VHSV tank experiments with herring under warmer ($\gamma_y = 0.192$ and $\alpha_y = 0.025$ day^{-1}) and cooler conditions ($\gamma_y = 0.055$ and $\alpha_y = 0.058$ day^{-1}; Table A1; Hershberger et al. 2013).

Estimation performance

Performance metrics were used to quantify the error in the estimates of model outputs. We used relative error (RE) and median absolute relative error (MARE) to measure bias (by the sign of RE) and precision (by the magnitude of RE and MARE) between the true value x_i and estimated value $\hat{x}_i$ for the simulated replicate i of the output:

$$\text{RE} = (\hat{x}_i - x_i)/x_i \quad (35)$$

$$\text{MARE} = \text{median}(|\hat{x}_i - x_i|/x_i) \quad (36)$$

RE was calculated for the 50 years of estimated spawning biomass and recruitment across all scenarios. We considered estimates to

Table 3. Summary of performance metrics of the estimation model in each scenario.

(*a*)	a_{50}^{mix}	a_{95}^{mix}	μ_y	Year 50 SB	SB	Recruits	ω_y
	MARE	MARE	MARE	MARE	Median MARE	Median MARE	Median DEV
Ignore disease	—	—	—	0.18	0.08	0.16	—
Incorporate infection prevalence	—	—	—	0.16	0.07	0.14	—
Incorporate seroprevalence	0.02	0.03	0.07	0.07	0.05	0.10	0.03
Small sample size	0.02	0.07	0.16	0.10	0.07	0.22	0.06
Early mixing of susceptible fish	0.22	0.18	0.13	0.06	0.05	0.13	0.05
Age-specific mixing ignored	—	—	0.89	0.42	0.22	2.69	0.10
Time-varying background mortality	0.02	0.03	0.07	0.12	0.09	0.13	0.02
Time-varying background mortality and ignore disease	—	—	—	0.26	0.11	0.19	—
Time-varying disease mortality and recovery	0.02	0.03	—	0.07	0.05	0.10	0.02
(*b*)	$P_{>0.4}$	$P_{>0.4}$	$P_{<0.4}$	$P_{<0.4}$			
	Median	Year 50	Median	Year 50			
Ignore disease	0.00	0.05	0.00	0.00			
Incorporate infection prevalence	0.00	0.05	0.00	0.00			
Incorporate seroprevalence	0.00	0.00	0.00	0.00			
Small sample size	0.00	0.01	0.00	0.00			
Early mixing of susceptible fish	0.00	0.00	0.00	0.00			
Age-specific mixing ignored	0.07	0.00	0.00	0.57			
Time-varying background mortality	0.00	0.02	0.00	0.01			
Time-varying background mortality and ignore disease	0.00	0.25	0.00	0.00			
Time-varying disease mortality and recovery	0.00	0.00	0.00	0.00			
(*c*)	$P_{>0.1}$	$P_{>0.1}$	$P_{<0.1}$	$P_{<0.1}$			
	Median	Year 50	Median	Year 50			
Ignore disease	0.06	0.75	0.16	0.01			
Incorporate infection prevalence	0.05	0.69	0.13	0.01			
Incorporate seroprevalence	0.05	0.20	0.08	0.14			
Small sample size	0.13	0.24	0.19	0.27			
Early mixing of susceptible fish	0.05	0.16	0.08	0.14			
Age-specific mixing ignored	0.46	0.01	0.10	0.98			
Time-varying background mortality	0.19	0.33	0.23	0.23			
Time-varying background mortality and ignore disease	0.17	0.81	0.28	0.01			
Time-varying disease mortality and recovery	0.09	0.19	0.10	0.14			

Note: Values are not provided (—) for parameters that are not estimated in a select few scenarios. Median MARE and median DEV were taken across years 2–50. Terms are defined in Table A1.

have little bias when the proportion of RE greater than 0 across all replicates was within 40%–60% (i.e., near a 50/50 chance of positive or negative RE). We also calculated the error in estimates of annual infection rates (ω_y), when included in the estimation model; the true ω_y is the total new infections in a year divided by the number of susceptible fish at the start of the transmission period. Because ω_y is bounded between 0 and 1 and can be zero or close to zero in some years (see Results), annual error was summarized as a simple deviation (DEV) between the true and estimated values:

$$\text{DEV} = \hat{\omega}_i - \omega_i \quad (37)$$

MAREs were calculated for spawning biomass for the final year of the modelled period and for spawning biomass and recruitment as the median of all year-specific MAREs from years 2 through 50. The median DEV of ω_y was calculated over years 2 through 49 (the final year of estimation was excluded as seroprevalence was sampled prior to and did not inform the current year's infection). MAREs are also presented for the main disease-related parameter estimates: recovery probability (λ), age at 50% (a_{50}^{mix}), and age at 95% (a_{95}^{mix}) mixing with the reservoir population. The true λ was derived from the operating model as the daily recovery rate divided by the sum of the daily rates of recovery and mortality from disease ($\gamma/(\alpha + \gamma)$) and was constant in all years in most scenarios (see Table A1).

Population assessments aim at providing accurate estimates of stock status and trends, while minimizing the risk of large errors in estimation, here defined as the proportion of simulations where estimated spawning biomass ($\hat{\text{SB}}_{y,i}$) was either more than 40% above ($P_{>0.4}$) or less than 40% below ($P_{<0.4}$) the true spawning biomass $\text{SB}_{y,i}$. The threshold of 40% has been used previously to represent the risk of large estimation error (e.g., Punt et al. 2018) as a reflection of the average error demonstrated in data-rich stock assessments of various US West Coast groundfish and coastal pelagic species (Ralston et al. 2011). However, most scenarios shown very low probabilities (Table 3*b*), so we also report probabilities with a threshold of 10% above and below ($P_{>0.1}$ and $P_{<0.1}$) to better differentiate the relative risk (Table 3*c*). We report these probabilities for the final year of the modelled period, as well as the median of the probabilities calculated for each year.

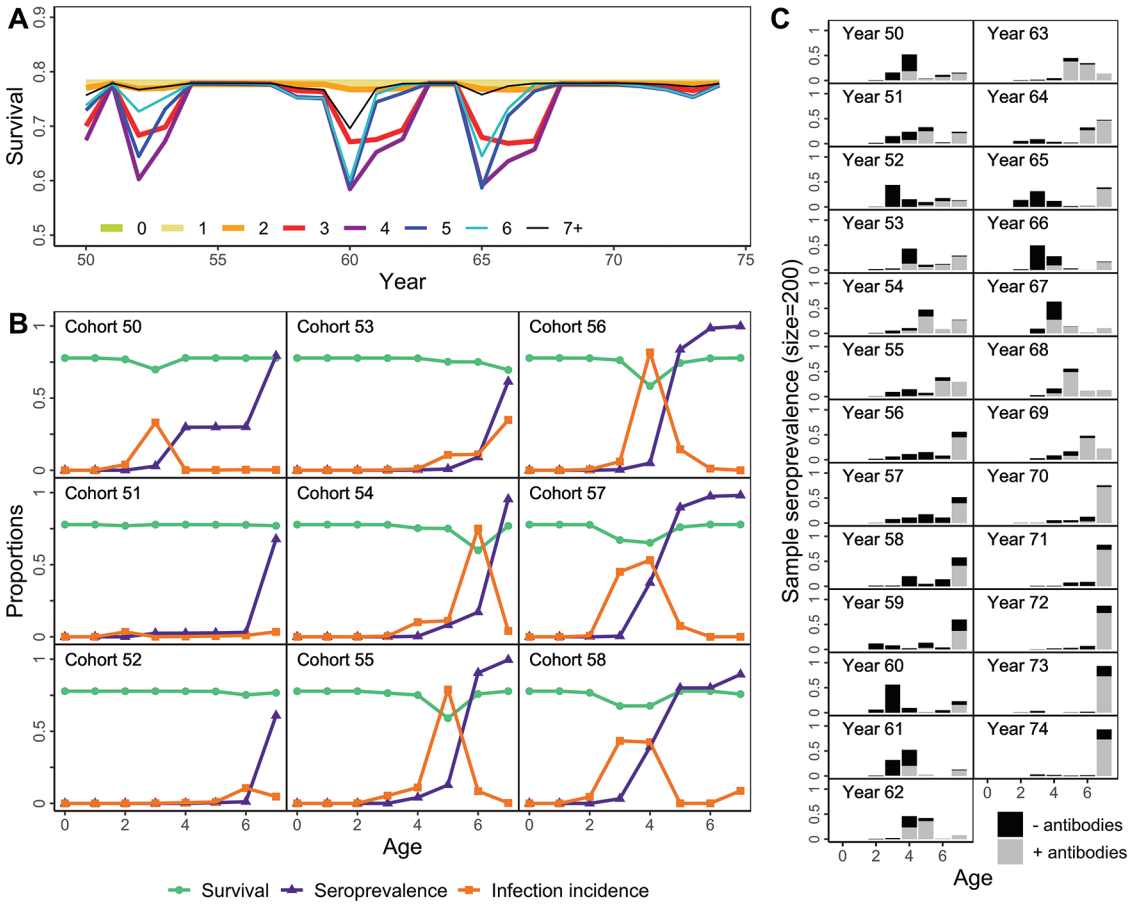
We checked for convergence of all estimation models fit in each scenario from the convergence codes output by the "nlminb" function within R (R Core Team 2020). Across all scenarios, 98%–100% of the estimation models converged. Results are presented only for converged models.

Results

Operating model characteristics

A 50-year subset of an example simulation of 100 years with stochastic recruitment showed highly dynamic interannual disease outbreaks (Fig. 2). Most years had low infection prevalence and were interspersed with years that have outbreaks (Figs. 2A–2C). The average time that an outbreak (infection prevalence $>$ 0.1)

Fig. 3. Age-specific time series of survival, seroprevalence, and infection in a subset of years. The upper left plot A displays age-specific survival over time in years 50 through 74 of the dynamics shown in Fig. 2. The lower left plot B displays age-specific profiles by cohort (i.e., Cohort 50 is the age 0 fish in year 50, age 1 in year 51, etc.) of survival, seroprevalence, and infection incidence experienced at each age. The right-most plot C is a simulated seroprevalence sample from the multinomial distribution using probabilities from the actual age-specific seroprevalence (accounting for survey selectivity) and a sample size of 200. Note the *y* axis in plot A starts at 0.5 and ends at 0.9. [Colour online.]



lasted was 24 days. Cross-correlations between annual population characteristics and disease-derived quantities from 500 years of simulation indicated recruitment lagged 3–4 from the past is positively correlated with current infection prevalence (ρ = 0.71–0.74); in other words, outbreaks occurred when large year classes fully mixed with the reservoir group (at ages 3–4). However, large differences in the amount of recruitment did not translate to large differences in the magnitude of infection prevalence or survival, suggesting that age and disease stage structure regulated the impacts of outbreaks (Fig. 2F).

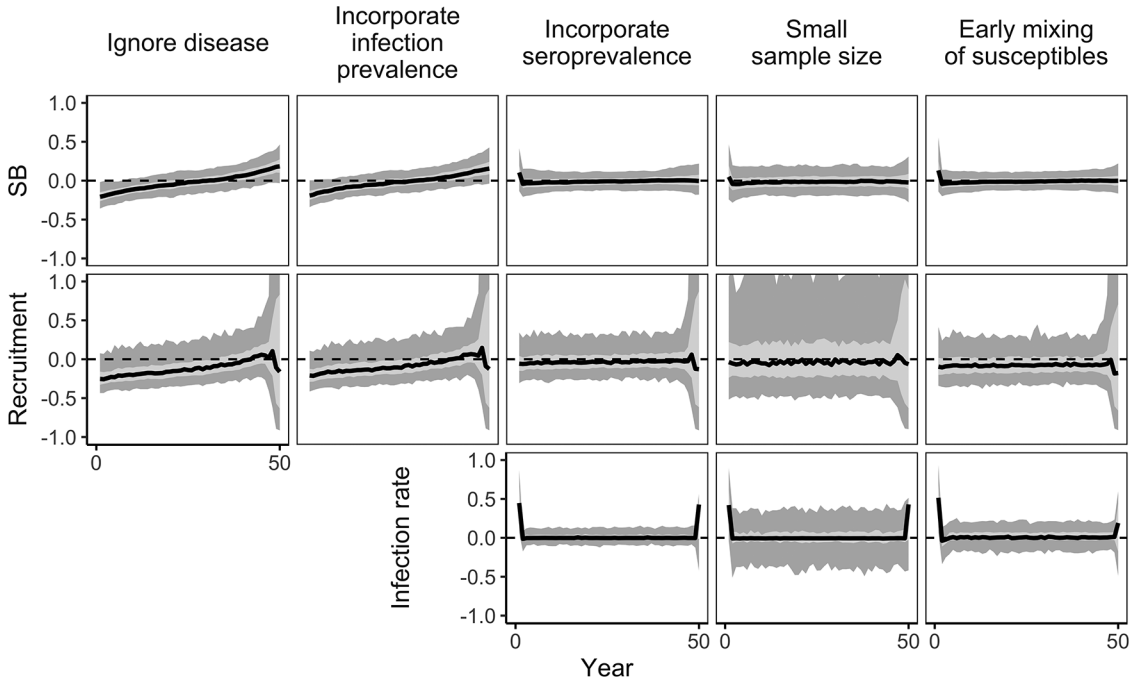
Seroprevalence increased in years that followed outbreaks (Fig. 2). Seroprevalence is measured prior to the transmission period in each year and reflects outbreaks in previous years. With increasing time since the most recent outbreak, seroprevalence decreased as immune cohorts declined in abundance from natural mortality (Fig. 2F). Years with higher infection rates had relatively higher disease-associated mortality, but this was followed by virtually no disease-associated mortality in subsequent years because of acquired herd immunity following outbreaks.

Changes in the proportions of susceptibles and carriers also determined differences in age-specific survival, infection, and seroprevalence within the operating model (Fig. 3). Reductions in survival and increases in infection incidence occur no earlier than age 3 (across years and within cohorts), which reflects the effect of age-specific mixing with the reservoir population on age-specific immunity and mortality (Figs. 3A and 3B). Starting at age 4, immunity (seroprevalence) typically increases with age as a greater proportion of each cohort experienced infection in their lifetime (and at lower infection incidence the older the cohort became; Cohorts 54–58 in Fig. 3B). However, some cohorts did not experience infection and acquire immunity late in life as well (Cohorts 51–53). Thus, outbreaks were more likely to occur when large cohorts mixed with the reservoir population and when multiple older age groups (e.g., >age 4) had low immunity (Fig. 3C).

Estimation without seroprevalence data

Population estimates had an increasing trend in bias (e.g., >90% of spawning biomass RE was negative in early years and positive in recent years) when seroprevalence data were not included in the estimation model (Fig. 4). The bias in spawning biomass and recruitment from the "Ignore disease" scenario stemmed from the estimation model missing the time- and age-varying natural mortality due to additional disease-associated mortality (Fig. 4). The bias of these estimates was accentuated when the instantaneous rate of mortality from disease (0.082 day^{-1}) exceeded that of recovery (0.014 day^{-1}),

Fig. 4. Time trajectories of error in estimates of spawning biomass (SB), recruitment, and annual infection rate (ω_y) across 50 years of simulation from five scenarios (Table 2). Relative error (RE) is calculated between the actual values (from the operating model) and the estimated values (from the estimation model) for SB and recruitment. Deviation (DEV) is shown for infection rates because these values are scaled between 0 and 1. The black denotes the median relative error or error across converged simulations, the light gray denotes the inner-50th percentiles, and dark denotes inner-95th percentiles.



which provided daily transition probabilities of $\alpha = 0.08$ and $\gamma = 0.01$ (results not shown). As a result, the RE was greater when annual recovery rate of infected fish was low (e.g., $\mu = 0.14$) and mortality from disease was high.

A similar trend in estimation bias of spawning biomass and recruitment resulted when using the infection prevalence index to account for disease-associated mortality within the estimation model (Incorporate infection prevalence). The MAREs for these estimates from using the infection prevalence index and ignoring disease scenarios were also similar (Table 3). Thus, unless surveys happen to coincide with peak of an epizootic (which is unlikely in our simulations), using an infection prevalence index was akin to disregarding the impact of disease on natural mortality.

Estimation with seroprevalence data

The estimation model with seroprevalence data (Incorporate seroprevalence) produced little bias in estimates (e.g., 43% of spawning biomass RE > 0 averaged over years; Fig. 4). Lower MAREs (Table 3*a*) of annual spawning biomass and recruitment are also shown compared to the scenarios without seroprevalence, and accurately estimated annual infection, age-specific mixing, and disease recovery rate. Estimates of spawning biomass in the final year improved the most (48% of the RE > 0 and low MAREs in Table 3*a*). In contrast, recruitment estimates in the final year had large RE in either direction (>0.7 or < –0.7; Fig. 4). This occurred because cohorts gradually enter the spawning population as they age and assessment models require multiple years of data on a single cohort to accurately estimate the total abundance of a single cohort (i.e., when the entire cohort is in the spawning population).

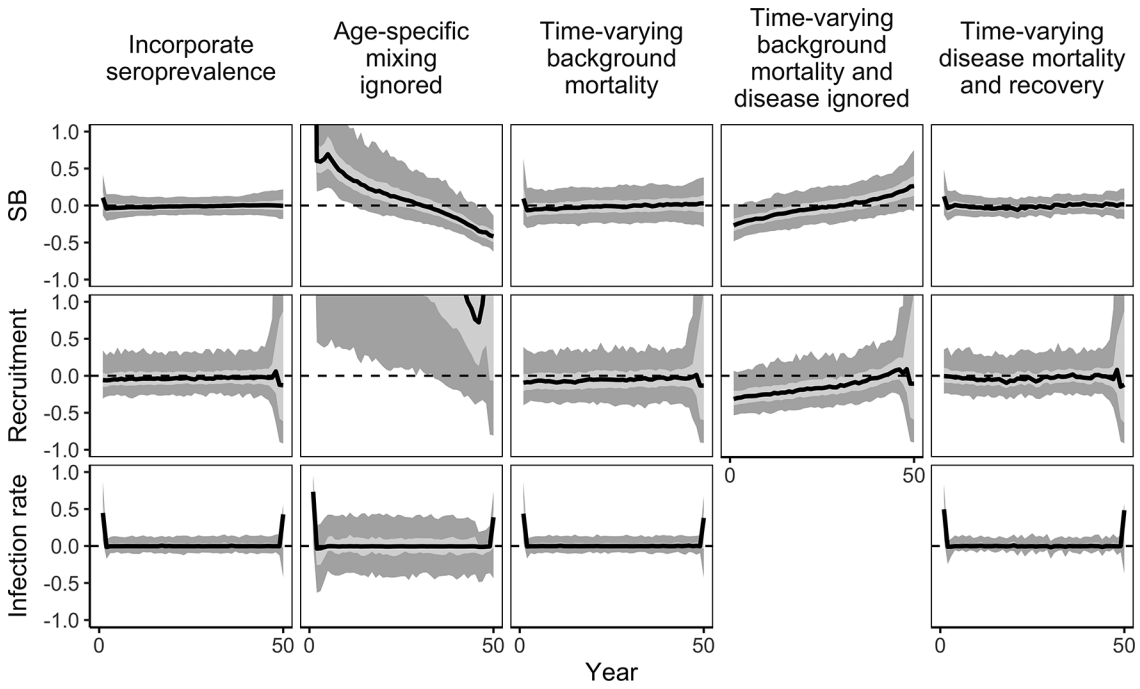
Estimation with fewer seroprevalence data or infection in unobserved ages

The estimation model still provided little bias in estimates of the spawning biomass and infection rates when seroprevalence sample sizes were small (Small sample size) or fish became vulnerable at ages earlier than when they were first observed in spawner surveys (Early mixing of susceptible fish; 42%–44% of spawning biomass RE > 0; Fig. 4). Smaller sample sizes for age composition and seroprevalence data resulted in less precise population and infection rate estimates compared to the equivalent scenario with large sample sizes (Incorporate seroprevalence; Table 3*a*) and increased the size of the error percentiles (Fig. 4). The model was still able to estimate the age-specific mixing with high precision, but the MARE around the recovery rate parameter was greater than in the "Incorporate seroprevalence" scenario (Table 3*a*). The MARE in these parameters was even higher in the scenario with early mixing of susceptible fish. Still, the estimation model in these two scenarios provided fairly accurate pictures of annual infection and its overall impact on the population.

Estimation with misspecifying age-specific mixing

If the estimation model assumed that fish of all ages are mixing in the reservoir population, and the operating model did not have this assumption (scenario "Age-specific mixing ignored"), then estimates of spawning biomass (69% of RE > 0 on average) and recruitment (98% of RE > 0 on average) had the largest bias amongst all scenarios (Fig. 5). The likely reason stems from assuming all young fish are immediately susceptible. Mortality from disease is then assumed to occur at the youngest ages and substantially greater recruitment must be estimated to compensate for this additional loss of young fish. The estimated recovery rate also has a large MARE (Table 3*a*). Altogether, this resulted in consistent underestimates of spawning biomass in recent years (Fig. 5). Annual infection still had little bias despite having larger RE (Fig. 5). Thus, it is crucial to carefully consider the ages at which fish first mix with (or at least become exposed to transmission from) reservoir populations.

Fig. 5. Time trajectories of error in estimates of spawning biomass (SB), recruitment, and annual infection rate (ω_y) across 50 years of simulation from the Incorporate seroprevalence scenario and the remaining four scenarios (Table 2). Presentation as in Fig. 4.



Estimation with time-varying rates

Ignoring disease and assuming constant background natural mortality while true background mortality varied (scenario "Time-varying background mortality and ignore disease") produced similar bias patterns to "Ignoring disease", but with even higher MAREs in spawning biomass and recruitment (Table 3*a*). When seroprevalence data were included in an estimation model with constant background mortality and time-varying mortality in the operating model (scenario "Time-varying background mortality"), mostly little bias in estimates of spawning biomass resulted (46% of RE $>$ 0 on average), but recruitment was slightly underestimated (62% of RE $<$ 0 on average; Fig. 5). This is likely because annual variation in the true background mortality (M_y) fluctuated lognormally around the same mean assumed fixed in the estimation model. In other words, the estimation model averages out annual deviations in background mortality, but with higher MAREs in spawning biomass and recruitment (Table 3*a*). Parameter estimates related to disease (annual infection, recovery rate, and age-specific mixing) also show low MARE (Table 3*a*). Thus, including seroprevalence provided improvement even when other influences on natural mortality were misspecified.

The estimation model performed well when the true recovery rate varied annually (μ = 0.52–0.87 over 100 years) but was assumed constant over time in the estimation model (scenario "Time-varying disease mortality and recovery"). Population and infection estimates had little bias across most years (45% of spawning biomass and 42% of recruitment RE $>$ 0 on average; Fig. 5), and the MAREs of these estimates and disease-related parameters were low and similar to the more simplistic "Incorporate seroprevalence" scenario (Table 3*a*). When the range of the variation in recovery rates was shifted lower (μ = 0.09–0.44), causing larger mortality from disease and more pronounced annual variation in total survival in the operating model, this resulted in higher REs and MAREs similar to scenario "Ignore disease" (results not shown). Thus, ignoring time variation in recovery rates when they are low in most years (e.g., $<$0.5) degrades performance of the estimation model.

Risk of substantial bias in estimates under different scenarios

A more practical consideration of the performance of estimation for stock assessment is to evaluate the risk of substantial overestimation or underestimation of stock status (the metrics $P_{>0.4}$, $P_{<0.4}$, $P_{>0.1}$, and $P_{<0.1}$). Scenarios that incorporated seroprevalence data for estimation show little risk that spawning biomass estimates were substantially wrong (Table 3*b*). Scenarios with time-varying rates (background or mortality from disease) also had little risk of substantial overestimation or underestimation. In contrast, ignoring age-specific mixing in estimation models resulted in high risk of severe underestimation of the final year biomass, while ignoring seroprevalence data led to positively biased spawning biomass estimates (Table 3*b* and 3*c*). These differences in risk between scenarios are better illustrated when considering a 10% threshold (Table 3*c*). Generally, including seroprevalence data to account for disease-associated mortality lowered the risk of substantial over- and underestimation.

Discussion

This study is the first proof-of-concept that seroprevalence data can be used to estimate mortality from disease within a fish stock assessment model. We provide a modelling framework that integrates epidemiological principles within fish population dynamics models to provide a realistic picture of disease impacts and improve estimation for fisheries management.

There are two key benefits to understanding exposure history to pathogens like VHSV. First, the recovery rate of infected fish (μ) can be directly estimated in the estimation model. This is possible for two reasons: (1) the survivors are identified and changes in abundance and age structure due to disease (and recruitment) are observed; and (2) the disease progression is unidirectional (i.e., infected individuals can become carriers or die). As a result, once the epizootic is over, the proportion of infected individuals that die is one minus the recovery rate. Second, knowing the current seroprevalence in the population also provides an understanding of the risk of increased mortality from disease in the

near future because high seroprevalence indicates widespread immunity and a low potential for future disease outbreaks.

Estimation of spawning biomass and recruitment by our model depends on model assumptions and can show potentially large errors (Figs. 4 and 5). Generally, scenarios where seroprevalence was incorporated and mortality from disease was modeled performed better than those without; however, inappropriate model assumptions about background mortality, disease recovery and mortality, or age-specific mixing can counteract improvements to estimation. Characteristics of the disease itself also determine the amount of improvement offered by incorporating seroprevalence; ignoring deadlier diseases (low recovery and high mortality from disease), or diseases with highly variable mortality rates can lead to large estimation errors.

Our estimation model correctly specified background mortality (M_y). Assuming background mortality is higher or lower than the true mortality would bias mean recruitment and disease recovery estimates (μ). For example, using higher background mortality in the Prince William Sound stock assessment resulted in larger estimates of recruitment because it could still explain the biomass levels observed on surveys (Muradian et al. 2017). Estimates of disease recovery (and thus mortality) would also scale with estimates of background mortality to explain observed biomass. In contrast estimates of annual infection should not be biased by the model because these are informed by seroprevalence data. However, bias in annual infection could arise from potential biases in the collection and calculation of seroprevalence data, which should be accounted for prior to including these data in the model.

We introduce the concept of age-specific mixing with the reservoir group (represented by v_a) and found that ignoring age-specific mixing when present was especially consequential for population estimates compared to most other scenarios. This occurs because age-specific mixing had a substantial impact on the characteristics and patterns of outbreaks (e.g., the ages most impacted by the extent and timing of outbreaks). For example, re-running all scenarios without age-specific mixing in both operating and estimation models ($v_a = 1$) resulted in all outbreaks occurring at age 0 before the fish are observed in surveys (not shown). This allowed all scenarios to have mostly little bias in estimates of spawning biomass because the model attributed the lower abundance of spawning fish (due to mortality from disease) to lower numbers of age 0 fish being spawned and recruited. In other words, if mortality from disease primarily occurs in the unobserved ages, the model cannot estimate the total impact of disease on these groups, but it can still estimate the overall population.

Our concept of age-specific mixing, and thus age-specific exposure to disease, captures the more complex transmission mechanics underlying diseases in fish populations that are not homogeneous. Age-specific mixing may be most evident in salmon species that have distinct life stages, such as young fish recruiting to an adult reservoir group when disease may act as a mechanism for overcompensation. There may be other reasons for age-specific mixing or different interpretations for age-specific vulnerability to transmission, such as the proportion of fish in an age class that congregate in embayments where conditions may be more conducive for epizootics (Hershberger et al. 2016). Thus, our parameterization for v_a (the proportion of juveniles mixing with adult fish) can be modified to have a more generic categorical representation for differences in transmission across age groups. For example, v_a could act as a multiplier to scale the infection rate between juvenile and adult fish if each group follows different exposure trajectories. Then v_a can in theory be larger for younger ages than older ages, or follow a different function other than the logistic (e.g., a block function, with fixed values for different groups of ages). Future research should focus on refining this concept for modeling the impact of age-structured mixing on epidemiology in stock assessments and evaluate alternative parameterizations for v_a or even different models to account for age-specific differences in transmission and thus infection.

VHSV in Pacific herring provided an ideal case study for our simulations because of extensive disease monitoring in this species over the last three decades (Elston and Meyers 2009; Hershberger et al. 1999; Marty et al. 1998, 2003, 2010; Meyers et al. 1994). Disease monitoring evolved to improve the VHSV information fed to the Prince William Sound herring stock assessment to move beyond the use of infection prevalence indices (Marty et al. 2003, 2010; Quinn et al. 2001), which have been criticized (Elston and Meyers 2009; Hershberger et al. 2016). Our simulation results showed that using infection prevalence performed just as badly as ignoring disease-associated mortality. This is a consequence of infection prevalence being measured infrequently (in our simulations and for Prince William Sound herring) and the inappropriate assumption that prevalence scales with mortality from disease. Infection prevalence may still be useful if regularly measured, or even infrequently measured if the diseases are slow progressing or endemic, when combined with appropriate models (e.g., multi-state mark–recapture models). Altogether, our results emphasize that disease recognition within population models should be carefully tailored to the characteristics of the host–pathogen interaction, the sampling protocol, and the model structure.

More generally, antibodies can be thought of as "natural" markers or tags in mark–recapture studies, much like the use of DNA in close-kin mark–recapture (Bravington et al. 2016). Multistate mark–recapture analyses of tagging data with infection prevalence (e.g., Groner et al. 2018) and the equations of our estimation model do not make any assumptions about the underlying transmission function (e.g., frequency dependence). The use of antibody "tags" detected with serological tests avoids some issues with capture methods in mark–recapture studies, such as tagging mortality, but retain the challenges in recapturing "marked" fish, especially collecting large samples that span all ages and represent a fully mixed population. Both approaches are challenged by spatial heterogeneity in disease transmission and outbreaks. Mark–recapture analyses allow for the estimation of the rates of disease stage transitions (e.g., susceptible to infected), while our estimation model does not, making these analyses more appropriate for informing the underlying epidemiology of these diseases. Still, our estimation model provides an alternative to mark–recapture analyses in estimating annual population-level infection and mortality from disease. Our approach also integrates seroprevalence data with other population survey data within a stock assessment model, thus directly informing stock status for fisheries management, where previous studies obtained estimates of mortality from disease in fish populations with mark–recapture models outside of stock assessment (Groner et al. 2018; Hoenig et al. 2017).

Accounting for age- and time-varying natural mortality remains a key challenge for stock assessment, but an increasingly understood topic due to recent simulation studies and reviews (Deroba and Schueller 2013; Jiao et al. 2012; Johnson et al. 2015; Lee et al. 2011; Punt et al. 2021). For example, effects of age-varying natural mortality are secondary to time-varying mortality on model performance (Deroba and Schueller 2013; Miller and Hyun 2018), but ignoring age variation can substantially bias model estimates under certain conditions, especially depending on how age-specific selectivity is modeled (He et al. 2011). These studies focus on freely estimating varying natural mortality (i.e., without ancillary data). Other studies have more explicitly attributed variation in mortality to predation (Trijoulet et al. 2020), mass die-offs due to environmental phenomena (e.g., red tides; Harford et al. 2018), or biological theory (e.g., as it relates to size and growth, or the "U-shape" assumption across ages; Chu et al. 2008; Gislason et al. 2010; Lorenzen 1996). Impacts of infectious diseases have been largely ignored in this body of literature. We show disease induces both age- and time-varying mortality (Fig. 3) and ignoring this degrades population estimates, as found in other studies that have investigated process errors in natural mortality (Miller and Hyun 2018; Trijoulet

et al. 2020). Thus, our work opens the door to explicitly accounting for an important source of variation in mortality in single-species stock assessments.

Fishing and its effects are usually incorporated into and estimated by models used for stock assessment, and comprise a component of total mortality we ignored in this study. Incorporating fishing mortality may impact the estimation performance of our model with seroprevalence depending on data quality, contrast in the observed dynamics, model assumptions, confounding process variation, and life history of the assessed population. We anticipate that accounting for fishing using consistent fishery-dependent data with high contrast over time (Lee et al. 2011) and the current assumptions of our estimation model will result in similar estimates of infection and disease mortality. We base this expectation on our estimation model having perfect information for commonly uncertain quantities (e.g., the relative abundance survey scalar q, effective sample sizes for composition and seroprevalence data) and simplified functions (e.g., maturity matching survey selectivity, constant growth). Realistically, such information and conditions do not exist for most assessed populations, which means mortality and fishing-related parameters, such as dome-shaped selectivity, may not be identifiable if they are estimated (He et al. 2011). Selectivity is particularly problematic because it's also likely time-varying and confounding with other time-varying processes (i.e., an underdetermined problem; Butterworth and Punt 1990). Estimation bias may increase with unknown effective sample sizes for age composition and seroprevalence (e.g., if the true effective sample size is much lower than assumed, as suspected from nonrandom schooling of fish), or be greater for other quantities, such as spawning biomass compared to recruitment if age- and time-varying mortality is estimated in long-lived species (Deroba and Schueller 2013). These known issues in stock assessment should be simulation tested in future work, but the complex and often case-specific scenarios in which they occur will require models tailored to the particular study population and pathogen for which antibody data are available.

We showed that ignoring disease effects, and thus misspecifying natural mortality, may lead to overestimation of the spawning biomass in the final year, which is used to determine current stock status or as a starting point for forecasts used for management decisions. Even with the relatively small changes in survival produced by our operating model, consistent overestimates in biomass result in a systemic positive bias of stock status that risks decisions leading to overexploitation. While we did not conduct the reference point estimation (e.g., B_{MSY} and F_{MSY}, the biomass and fishing mortality at maximum sustainable yield) also needed to determine stock status, errors in natural mortality from misspecification or estimation have been shown to be directly related to errors in reference points (Punt et al. 2021). For example, overestimating natural mortality caused a positive bias in the target fishing mortality rate, but the impact of errors in natural mortality can be reduced by feedback management strategies (Punt et al. 2021). Impacts to fisheries management may further be limited if relative reference points are used (e.g., B_{MSY} and current B are biased, but not their ratio). The potential benefit of our proposed method for fisheries management needs further investigation, whether that is determining how seroprevalence data may improve reference point estimation, or testing management strategies that are robust to disease mortality and the estimation error from misspecifying it.

Our operating and estimation models can be adapted to and should be tested with other directly transmitted pathogens and host populations. In the operating model, parameter value changes such as the transition probabilities within the daily disease transition matrix (eq. 5) may be changed to reflect host (e.g., contact or encounter frequency) and pathogen (e.g., infectiousness) characteristics. These can be obtained from laboratory experiments or regularly measured prevalence in wild populations after selectivity due to disease is accounted for. The number and types of stages could also be modified to reflect variation in disease progression (e.g., the existence of a latent infection stage, or the absence of a recovered stage). Generalized forms of transmission function have been developed by including powers to allow scaling between density- and frequency-dependent factors (Hopkins et al. 2020). Recruitment, the only stochastic component in our operating model, was assumed to occur before transmission starts; however, the fine-scale timing of recruitment (whether age 0 fish or older fish joining spawning aggregations) within a year can be modeled and has been identified as a crucial determinant of the timing of infection and pathogen persistence in wildlife populations (Begon et al. 2009; Peel et al. 2014; Smith et al. 2008). Still, the epidemiological equations in our estimation model circumvent many assumptions regarding transmission because the estimated annual infection rates reflect the emergent dynamics, not the underlying mechanics of the host–pathogen interaction. Therefore, the forms of eqs. 27–29 should be more generally applicable to population models of other species, although different life histories may impact estimation performance. The one exception to this general applicability is the proportion of fish that remain infected may have to be explicitly accounted for diseases with substantial numbers of infected individuals surviving between years (i.e., modifying eqs. 27 and 28).

Models of seroprevalence in wildlife populations are a fairly recent development. Such models have been used to quantify exposure of wild ruminant species to Schmallenberg virus in France (Rossi et al. 2017), measure the impact of detection errors on predicting seroprevalence of viruses in wild pigs in the US (Tabak et al. 2019), and identify environmental risk factors in seroprevalence patterns of an endemic tick-borne pathogen in sheep as transmitted from red grouse and deer in Scotland (Gilbert et al. 2020). Our study adds a novel and unique contribution to this emerging field within ecology as a means to integrate seroprevalence data with demographic data to jointly estimate population-level infection history and population status. For the field of fisheries science, we provide a more rigorous way to model age- and time-varying survival due to disease to improve fish stock assessments of affected populations, which can be extended to other pathogen types. Our modeling framework provides a major advance to the toolset needed by fisheries management to account for mortality from disease in fish populations.

Acknowledgements

Funding for this work was provided by the Exxon Valdez Oil Trustee Council (EVOSTC). Trevor Branch was partially funded by the Richard C. and Lois M. Worthington Endowed Professor in Fisheries Management. We thank A. Punt (University of Washington), T. Essington (University of Washington), O. Shelton (NOAA), R. Perry (USGS), and two anonymous reviewers for invaluable comments on the manuscript. Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the US Government. The authors declare there are no competing interests.

References

Anderson, R.M., and May, R.M. 1979. Population biology of infectious diseases: Part I. Nature, **280**(5721): 361–367. doi:10.1038/280361a0. PMID:460412.

Begon, M., Telfer, S., Smith, M.J., Burthe, S., Paterson, S., and Lambin, X. 2009. Seasonal host dynamics drive the timing of recurrent epidemics in a wildlife population. Proc. Biol. Sci. **276**(1662): 1603–1610. doi:10.1098/rspb.2008.1732. PMID:19203924.

Bravington, M.V., Skaug, H.J., and Anderson, E.C. 2016. Close-kin mark–recapture. Stat. Sci. **31**(2): 259–274. doi:10.1214/16-STS552.

Briggs, C.J., Vredenburg, V.T., Knapp, R.A., and Rachowicz, L.J. 2005. Investigating the population-level effects of chytridiomycosis: an emerging infectious disease of amphibians. Ecology, **86**(12): 3149–3159. doi:10.1890/04-1428.

Butterworth, D., and Punt, A. 1990. Some preliminary examinations of the potential information content of age-structure data from Antarctic minke whale research catches. Rep. Int. Whaling Commission, **40**: 301–315.

Chu, C.Y.C., Chien, H.-K., and Lee, R.D. 2008. Explaining the optimality of U-shaped age-specific mortality. Theor. Popul. Biol. **73**(2): 171–180. doi:10.1016/j.tpb.2007.11.005. PMID:18178233.

Deroba, J.J., and Schueller, A.M. 2013. Performance of stock assessments with misspecified age- and time-varying natural mortality. Fish Res. **146**: 27–40. doi:10.1016/j.fishres.2013.03.015.

Elston, R.A., and Meyers, T.R. 2009. Effect of viral hemorrhagic septicemia virus on Pacific herring in Prince William Sound, Alaska, from 1989 to 2005. Dis. Aquat. Org. **83**(3): 223–246. doi:10.3354/dao02005.

Ford, S.E., and Haskin, H.H. 1982. History and epizootiology of *Haplosporidium nelsoni* (MSX), an oyster pathogen in Delaware Bay, 1957–1980. J. Invertebr. Pathol. **40**(1): 118–141. doi:10.1016/0022-2011(82)90043-X.

Garver, K.A., Traxler, G.S., Hawley, L.M., Richard, J., Ross, J., and Lovy, J. 2013. Molecular epidemiology of viral haemorrhagic septicaemia virus (VHSV) in British Columbia, Canada, reveals transmission from wild to farmed fish. Dis. Aquat. Organ. **104**(2): 93–104. doi:10.3354/dao02588. PMID:23709462.

Gaughan, D., Mitchell, R., and Blight, S. 2000. Impact of mortality, possibly due to herpesvirus, on pilchard *Sardinops sagax* stocks along the south coast of Western Australia in 1998–99. Mar. Freshwater Res. **51**(6): 601–612. doi:10.1071/MF99176.

Gilbert, L., Brülisauer, F., Willoughby, K., and Cousens, C. 2020. Identifying environmental risk factors for louping ill virus seroprevalence in sheep and the potential to inform wildlife management policy. Front. Vet. Sci. **7**: 377. doi:10.3389/fvets.2020.00377. PMID:32695800.

Gislason, H., Daan, N., Rice, J.C., and Pope, J.G. 2010. Size, growth, temperature and the natural mortality of marine fish. Fish Fish. **11**(2): 149–158. doi:10.1111/j.1467-2979.2009.00350.x.

Groner, M.L., Hoenig, J.M., Pradel, R., Choquet, R., Vogelbein, W.K., Gauthier, D.T., and Friedrichs, M.A.M. 2018. Dermal mycobacteriosis and warming sea surface temperatures are associated with elevated mortality of striped bass in Chesapeake Bay. Ecol. Evol. **8**(18): 9384–9397. doi:10.1002/ece3.4462. PMID:30377509.

Harford, W.J., Grüss, A., Schirripa, M.J., Sagarese, S.R., Bryan, M., and Karnauskas, M. 2018. Handle with care: establishing catch limits for fish stocks experiencing episodic natural mortality events. Fisheries, **43**(10): 463–471. doi:10.1002/fsh.10131.

Hart, L.M., MacKenzie, A., Purcell, M.K., Powers, R.L., and Hershberger, P.K. 2017. Optimization of a plaque neutralization test (PNT) to identify the exposure history of Pacific herring to viral hemorrhagic septicemia virus (VHSV). J. Aquat. Anim. Health, **29**(2): 74–82. doi:10.1080/08997659.2017.1285369. PMID:28375717.

Hay, D.E., Toresen, R., Stephenson, R., Thompson, M., Claytor, R., Funk, F., et al. 2001. Taking stock: an inventory and review of world herring stocks in 2000. *In* Proceedings of the 18th Lowell Wakefield Symposium: Herring Expectations For A New Millenium, Anchorage, Alaska, 23–26 February 2000. University of Alaska Sea Grant College Program, AK-SG-01-04. pp. 381–454.

Haydon, D.T., Cleaveland, S., Taylor, L.H., and Laurenson, M.K. 2002. Identifying reservoirs of infection: a conceptual and practical challenge. Emerg. Infect. Dis. **8**(12): 1468–1473. doi:10.3201/eid0812.010317. PMID:12498665.

He, X., Ralston, S., and MacCall, A.D. 2011. Interactions of age-dependent mortality and selectivity functions in age-based stock assessment models. Fish Bull. **109**(2): 198–216.

Heisey, D.M., Joly, D.O., and Messier, F. 2006. The fitting of general Force-of-infection models to wildlife disease prevalence data. Ecology, **87**(9): 2356–2365. doi:10.1890/0012-9658(2006)87[2356:TFOGFM]2.0.CO;2. PMID:16995636.

Hershberger, P.K., Kocan, R.M., Elder, N.E., Meyers, T.R., and Winton, J.R. 1999. Epizootiology of viral hemorrhagic septicemia virus in Pacific herring from the spawn-on-kelp fishery in Prince William Sound, Alaska, USA. Dis. Aquat. Org. **37**(1): 23–31. doi:10.3354/dao037023.

Hershberger, P.K., Gregg, J.L., Grady, C.A., Taylor, L., and Winton, J.R. 2010. Chronic and persistent viral hemorrhagic septicemia virus infections in Pacific herring. Dis. Aquat. Org. **93**(1): 43–49. doi:10.3354/dao02283.

Hershberger, P.K., Purcell, M.K., Hart, L.M., Gregg, J.L., Thompson, R.L., Garver, K.A., and Winton, J.R. 2013. Influence of temperature on viral hemorrhagic septicemia (Genogroup IVa) in Pacific herring, *Clupea pallasii* Valenciennes. J. Exp. Mar. Biol. Ecol. **444**: 81–86. doi:10.1016/j.jembe.2013.03.006.

Hershberger, P.K., Garver, K.A., and Winton, J.R. 2016. Principles underlying the epizootiology of viral hemorrhagic septicemia in Pacific herring and other fishes throughout the North Pacific Ocean. Can. J. Fish. Aquat. Sci. **73**(5): 853–859. doi:10.1139/cjfas-2015-0417.

Hoenig, J.M., Groner, M.L., Smith, M.W., Vogelbein, W.K., Taylor, D.M., Landers, D.F., Jr., et al. 2017. Impact of disease on the survival of three commercially fished species. Ecol. Appl. **27**(7): 2116–2127. doi:10.1002/eap.1595. PMID:28675580.

Hopkins, S.R., Fleming-Davies, A.E., Belden, L.K., and Wojdak, J.M. 2020. Systematic review of modelling assumptions and empirical evidence: Does parasite transmission increase nonlinearly with host density? Methods Ecol. Evol. **11**(4): 476–486. doi:10.1111/2041-210X.13361.

Jiao, Y., Smith, E.P., O'Reilly, R., and Orth, D.J. 2012. Modelling non-stationary natural mortality in catch-at-age models. ICES J. Mar. Sci. **69**(1): 105–118. doi:10.1093/icesjms/fsr184.

Johnson, K.F., Monnahan, C.C., McGilliard, C.R., Vert-Pre, K.A., Anderson, S.C., Cunningham, C.J., et al. 2015. Time-varying natural mortality in fisheries stock assessment models: identifying a default approach. ICES J. Mar. Sci. **72**(1): 137–150. doi:10.1093/icesjms/fsu055.

Klepac, P., and Caswell, H. 2011. The stage-structured epidemic: linking disease and demography with a multi-state matrix approach model. Theor. Ecol. **4**(3): 301–319. doi:10.1007/s12080-010-0079-8.

Kristensen, K., Nielsen, A., Berg, C., Skaug, H., and Bell, B. 2016. TMB: automatic differentiation and Laplace approximation. J. Stat. Soft. **70**(5): 1–21. doi:10.18637/jss.v070.i05.

Lee, H.-H., Maunder, M.N., Piner, K.R., and Methot, R.D. 2011. Estimating natural mortality within a fisheries stock assessment model: an evaluation using simulation analysis based on twelve stock assessments. Fish Res. **109**(1): 89–94. doi:10.1016/j.fishres.2011.01.021.

Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. J. Fish Biol. **49**(4): 627–642. doi:10.1111/j.1095-8649.1996.tb00060.x.

MacCall, A.D. 1990. Dynamic geography of marine fish populations. Washington Sea Grant Program, Seattle, Wash.

Marty, G.D., Freiberg, E.F., Meyers, T.R., Wilcock, J., Farver, T.B., and Hinton, D.E. 1998. Viral hemorrhagic septicemia virus, *Ichthyophonus hoferi*, and other causes of morbidity in Pacific herring *Clupea pallasi* spawning in Prince William Sound, Alaska, USA. Dis. Aquat. Org. **32**(1): 15–40. doi:10.3354/dao032015. PMID:9676259.

Marty, G.D., Ii, T.J.Q., Carpenter, G., Meyers, T.R., and Willits, N.H. 2003. Role of disease in abundance of a Pacific herring (*Clupea pallasi*) population. Can. J. Fish. Aquat. Sci. **60**(10): 1258–1265. doi:10.1139/f03-109.

Marty, G.D., Hulson, P.-J.F., Miller, S.E., Quinn, T.J., II, Moffitt, S.D., and Merizon, R.A. 2010. Failure of population recovery in relation to disease in Pacific herring. Dis. Aquat. Org. **90**(1): 1–14. doi:10.3354/dao02210.

Metcalf, C.J.E., Lessler, J., Klepac, P., Morice, A., Grenfell, B.T., and Bjørnstad, O.N. 2012. Structured models of infectious disease: Inference with discrete data. Theor. Popul. Biol. **82**(4): 275–282. doi:10.1016/j.tpb.2011.12.001. PMID:22178687.

Meyers, T.R., and Winton, J.R. 1995. Viral hemorrhagic septicemia virus in North America. Annu. Rev. Fish Dis. **5**: 3–24. doi:10.1016/0959-8030(95)00002-X.

Meyers, T.R., Short, S., Upson, K., Batts, W.N., Winton, J.R., Wilcock, J., and Brown, E. 1994. Association of viral hemorrhagic septicemia virus with epizootic hemorrhages of the skin in Pacific herring *Clupea harengus pallasi* from Prince William Sound and Kodiak Island, Alaska, USA. Dis. Aquat. Org. **19**(1): 27–37. doi:10.3354/dao019027.

Miller, T.J., and Hyun, S.-Y. 2018. Evaluating evidence for alternative natural mortality and process error assumptions using a state-space, age-structured assessment model. Can. J. Fish. Aquat. Sci. **75**(5): 691–703. doi:10.1139/cjfas-2017-0035.

Muradian, M.L., Branch, T.A., Moffitt, S.D., and Hulson, P.-J.F. 2017. Bayesian stock assessment of Pacific herring in Prince William Sound, Alaska. PLoS One, **12**(2): e0172153. doi:10.1371/journal.pone.0172153. PMID:28222151.

Murray, A.G. 2009. Using simple models to review the application and implications of different approaches used to simulate transmission of pathogens among aquatic animals. Prev. Vet. Med. **88**(3): 167–177. doi:10.1016/j.prevetmed.2008.09.006. PMID:18930326.

Peel, A.J., Pulliam, J.R.C., Luis, A.D., Plowright, R.K., O'Shea, T.J., Hayman, D.T.S., et al. 2014. The effect of seasonal birth pulses on pathogen persistence in wild mammal populations. Proc. Biol. Sci. **281**(1786): 20132962. doi:10.1098/rspb.2013.2962. PMID:24827436.

Punt, A.E., Okamoto, D.K., MacCall, A.D., Shelton, A.O., Armitage, D.R., Cleary, J.S., et al. 2018. When are estimates of spawning stock biomass for small pelagic fishes improved by taking spatial structure into account? Fish Res. **206**: 65–78. doi:10.1016/j.fishres.2018.04.017.

Punt, A.E., Castillo-Jordán, C., Hamel, O.S., Cope, J.M., Maunder, M.N., and Ianelli, J.N. 2021. Consequences of error in natural mortality and its estimation in stock assessment models. Fish Res. **233**: 105759. doi:10.1016/j.fishres.2020.105759.

Quinn, T., Marty, G.D., Wilcock, J., and Willette, M. 2001. Disease and population assessment of Pacific herring in Prince William Sound, Alaska. *In* Herring: Expectations For a New Millennium, Anchorage, Alaska, 23–26 February 2000. *Edited by* F. Funk, J. Blackburn, D. Hay, A.J. Paul, and R. Stephanson. University of Alaska Sea Grant, Anchorage, Alaska. pp. 363–379.

R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.R-project.org/.

Ralston, S., Punt, A.E., Hamel, O.S., DeVore, J.D., and Conser, R.J. 2011. A meta-analytic approach to quantifying scientific uncertainty in stock assessments. Fish Bull. **109**(2): 217–232.

Rossi, S., Viarouge, C., Faure, E., Gilot-Fromont, E., Gache, K., Gibert, P., et al. 2017. Exposure of wildlife to the Schmallenberg virus in France (2011–2014): higher faster, stronger (than bluetongue)! Transbound. Emerg. Dis. **64**(2): 354–363. doi:10.1111/tbed.12371. PMID:25958882.

Skall, H.F., Olesen, N.J., and Mellergaard, S. 2005. Viral haemorrhagic septicaemia virus in marine fish and its implications for fish farming — a review. J. Fish Dis. **28**(9): 509–529. doi:10.1111/j.1365-2761.2005.00654.x.
Smith, M.J., White, A., Sherratt, J.A., Telfer, S., Begon, M., and Lambin, X. 2008. Disease effects on reproduction can cause population cycles in seasonal environments. J. Anim. Ecol. **77**(2): 378–389. doi:10.1111/j.1365-2656.2007.01328.x. PMID:18005128.
Szuwalski, C.S., Vert-Pre, K.A., Punt, A.E., Branch, T.A., and Hilborn, R. 2015. Examining common assumptions about recruitment: a meta-analysis of recruitment dynamics for worldwide marine fisheries. Fish Fish. **16**(4): 633–648. doi:10.1111/faf.12083.
Szuwalski, C.S., Britten, G.L., Licandeo, R., Amoroso, R.O., Hilborn, R., and Walters, C. 2019. Global forage fish recruitment dynamics: a comparison of methods, time-variation, and reverse causality. Fish Res. **214**: 56–64. doi:10.1016/j.fishres.2019.01.007.
Tabak, M.A., Pedersen, K., and Miller, R.S. 2019. Detection error influences both temporal seroprevalence predictions and risk factors associations in wildlife disease models. Ecol. Evol. **9**(18): 10404–10414. doi:10.1002/ece3.5558. PMID:31632645.
Trijoulet, V., Fay, G., and Miller, T.J. 2020. Performance of a state-space multispecies model: What are the consequences of ignoring predation and process errors in stock assessments? J. Appl. Ecol. **57**(1): 121–135. doi:10.1111/1365-2664.13515.
Trochta, J.T., and Branch, T.A. 2021. Applying Bayesian model selection to determine ecological covariates for recruitment and natural mortality in stock assessment. ICES J. Mar. Sci. **78**(8): 2875–2894. doi:10.1093/icesjms/fsab165.
Vogelbein, W.K., Hoenig, J., Gauthier, D., Smith, M., Sadler, P., Kator, H., and Rhodes, M. 2012. The role of mycobacteriosis in elevated natural mortality of Chesapeake Bay striped bass: developing better models for stock assessment and management: a final report. Virginia Institute of Marine Science, William & Mary, Gloucester Point, Va. doi:10.25773/v5-d39g-3056
Wilson, A., Goldberg, T., Marcquenski, S., Olson, W., Goetz, F., Hershberger, P., et al. 2014. Development and evaluation of a blocking enzyme-linked immunosorbent assay and virus neutralization assay to detect antibodies to viral hemorrhagic septicemia virus. Clin. Vaccine Immunol. **21**(3): 435–442. doi:10.1128/CVI.00675-13. PMID:24429071.
Winter, A.K., Martinez, M.E., Cutts, F.T., Moss, W.J., Ferrari, M.J., McKee, A., et al. 2018. Benefits and challenges in using seroprevalence data to inform models for measles and rubella elimination. J. Infect. Dis. **218**(3): 355–364. doi:10.1093/infdis/jiy137. PMID:29562334.

Appendix A

Derivation of immune proportions equations in the estimation model

The proportion of each age within the population that is immune to infection in year y is $\widetilde{\Omega}_{y,a}$, so the numbers immune at each age are $N_{\text{immune},y,a} = \widetilde{\Omega}_{y,a}N_{y,a}$, To project the numbers immune in the next year $y + 1$ ($N_{\text{immune},y+1,a+1}$), we must account for background survival, the proportion immune in the current year $\widetilde{\Omega}_{y,a}$, and the proportion that are susceptible and vulnerable, but became infected and recover ($v_a E_{y,a}\omega_y\mu$):

$$N_{\text{immune},y+1,a+1} = \exp(-M_y)\left(\widetilde{\Omega}_{y,a} + v_a E_{y,a}\omega_y\mu\right)N_{y,a}$$

The new immune proportions, $\widetilde{\Omega}_{y+1,a+1}$, are then

$$\widetilde{\Omega}_{y+1,a+1} = \frac{\exp(-M_y)\left(\widetilde{\Omega}_{y,a} + v_a E_{y,a}\omega_y\mu\right)N_{y,a}}{N_{y+1,a+1}}$$

where $N_{y+1,a+1}$ is projected from eq. 27:

$$N_{y+1,a+1} = \exp(-M_y)\left[1 - v_a^{\text{trans}} E_{y,a}\omega_y(1-\mu)\right]N_{y,a}$$

Substituting $N_{y+1,a+1}$ back into the equation for $\widetilde{\Omega}_{y+1,a+1}$ above gives

$$\widetilde{\Omega}_{y+1,a+1} = \frac{\exp(-M_y)\left(\widetilde{\Omega}_{y,a} + v_a E_{y,a}\omega_y\mu\right)N_{y,a}}{\exp(-M_y)\left[1 - v_a E_{y,a}\omega_y(1-\mu)\right]N_{y,a}}$$

where s_a and $N_{y,a}$ cancel out to provide eq. 28:

$$\widetilde{\Omega}_{y+1,a+1} = \frac{\widetilde{\Omega}_{y,a} + v_a E_{y,a}\omega_y\mu}{1 - v_a E_{y,a}\omega_y(1-\mu)}$$

For the plus age group, both numbers from one year younger than the starting age of the plus group as well as the plus group in the current year are projected into the next year following

$$N_{y+1,n_a} = \exp(-M_y)\left[1 - v_{n_a-1}E_{y,n_a-1}\omega_y(1-\mu)\right]N_{y,n_a-1} + \exp(-M_y)\left[1 - v_{n_a}E_{y,n_a}\omega_y(1-\mu)\right]N_{y,n_a}$$

Calculating the immune numbers in the plus-group is a similar equation, but accounts for current immune proportions:

$$N_{\text{immune},y+1,n_a} = \exp(-M_y)\left(\widetilde{\Omega}_{y,n_a-1} + v_{n_a-1}E_{y,n_a-1}\omega_y\mu\right)N_{y,n_a-1} + \exp(-M_y)\left(\widetilde{\Omega}_{y,n_a} + v_{n_a}E_{y,n_a}\omega_y\mu\right)N_{y,n_a}$$

The next year's immune proportion of the plus group is then

$$\widetilde{\Omega}_{y+1,n_a} = \frac{\exp(-M_y)\left(\widetilde{\Omega}_{y,n_a-1} + v_{n_a-1}E_{y,n_a-1}\omega_y\mu\right)N_{y,n_a-1} + \exp(-M_y)\left(\widetilde{\Omega}_{y,n_a} + v_{n_a}E_{y,n_a}\omega_y\mu\right)N_{y,n_a}}{N_{y+1,n_a}}$$

Table A1. Descriptions of indices, derived quantities, and parameters used in the operating and estimation models.

Notation	Description	Type	Default values or deriving equations
OM–EM shared			
d	Day	Index	1–120
a	Age	Index	0–7+
y	Year	Index	1–100
j	Age transmitting infection (eqs. 6 and 7)	Index	0–7+
p	Index for combined positive and negative antibody prevalence by age	Index	1–14
d_{end}	Index of last day in season of transmission	Index	120
n_a	Plus-group age	Index	7
n_y	Number of years fit in estimation model	Index	50
$N_{y,a}$	Number of fish of age a during year y	Derived quantity	OM: $\sum(T_{y,d=1,a}, I_{y,d=1,a}, C_{y,d=1,a})$ EM: eqs. 30 and 31
v_a	Proportion of fish mixing with reservoir population as a function of age a	Derived quantity	Eq. 2
m_a	Maturity ogive as a function of age	Derived quantity	Eq. 20 ($m_a = s_a$)
s_a	Fishery-independent survey selectivity as a function of age a	Derived quantity	Eq. 20
$\widetilde{\Omega}_{y,a}$	Proportion of age a fish that are immune in year y	Derived quantity	Eqs. 23 and 28
$\widetilde{\Omega}_y$	Total proportion of population that are immune	Derived quantity	$\sum_{a=1}^{n_a} \widetilde{\Omega}_{y,a}$
w_a	Weight-at-age (grams)	Parameter	70, 94, 115, 134, 150, 160, 165, 168
M_y	Annual background natural mortality rate (day^{-1})	Parameter	0.25
$\overline{R}$	Average age-0 recruitment (millions)	Parameter	181.27 (Est)
ε_y	Annual deviations in recruitment about constant average recruitment	Parameter	—
σ_R	Standard deviation of annual recruitment	Parameter	1.2 (Est)
P	Extent of autocorrelation in recruitment deviations	Parameter	0.6
a_{50}^{mix}	Age at which 50% are mixing with reservoir population	Parameter	3 or 1 (Est)
a_{95}^{mix}	Age at which 95% are mixing with reservoir population	Parameter	4 (Est)
a_{50}^{survey}	Age at which 50% are selected for fishery-independent sampling	Parameter	3 (Est)
a_{95}^{survey}	Age at which 95% are selected for fishery-independent sampling	Parameter	4 (Est)
q	Survey scalar	Parameter	0.606
σ_B	Standard deviation for lognormal errors on biomass survey indices	Parameter	0.3
OM specific			
$\mathbf{T}_{d,a}$	Disease stage transition matrix of age a in day d	Derived quantity	Eq. 5
$T_{y,d,a}$	Number of total susceptible individuals of age a in year y on day d	Derived quantity	Eq. 10
$S_{y,d,a}$	Number of reservoir susceptible individuals of age a in year y on day d	Derived quantity	Eq. 1
$I_{y,d,a}$	Number of active infected individuals of age a in year y on day d	Derived quantity	Eq. 11
$C_{y,d,a}$	Number of carrier individuals of age a in year y on day d	Derived quantity	Eq. 12
$\boldsymbol{n}(d)$	Vector of all numbers of age a fish in each disease stage in year y on day d	Derived quantity	Eq. 3
$\lambda_{I,d,a}$	Probability of becoming infected due to contacts with infected individuals	Derived quantity	Eq. 6
$\lambda_{C,d,a}$	Probability of becoming infected due to contacts with carrier individuals	Derived quantity	Eq. 7
SB_y	True spawning biomass of fish in year y	Derived quantity	Eq. 19
B_y	Simulated observed biomass of fish in year y	Derived quantity	Eq. 21
$\Theta_{y,a}$	Simulated age-composition data in fishery-independent sample in year y	Derived quantity	Eq. 22
$A_{y,a}^{+}$	Simulated positive antibody numbers at age a in year y	Derived quantity	Eq. 24
$A_{y,a}^{-}$	Simulated negative antibody numbers at age a in year y	Derived quantity	Eq. 25
$\boldsymbol{A_{y,p}}$	Combined simulated numbers with positive and negative antibodies across ages	Derived quantity	Eq. 26
$\Psi_{y,d}$	Infection prevalence within total population on day d of year y	Derived quantity	Eq. 33
$\overline{\Psi}_y$	Mean infection prevalence from simulated samples of true infection prevalence in year y	Derived quantity	$\sum_1^3 \Psi_{y,d}/3$
M_d	Daily background natural mortality rate (day^{-1})	Parameter	6.85×10^4

Table A1 (*concluded*).

Notation	Description	Type	Default values or deriving equations
α	Daily probability of death due to disease given infection (based on instantaneous rate 0.030 day^{-1} under ambient temperatures; Hershberger et al. 2013)	Parameter	0.0297
γ	Daily probability of recovery from disease given infection (based on instantaneous rate 0.071 day^{-1} under ambient temperatures; Hershberger et al. 2013)	Parameter	0.0688
β_I	Mean transmission rate of disease from infected individuals to susceptible (day^{-1})	Parameter	0.01
β_C	Mean transmission rate of disease from carrier individuals to susceptible (day^{-1})	Parameter	0.000001
α_y	Time-varying probability of death due to disease (based on rates 0.025–0.058 day^{-1})	Parameter	0.025–0.056
γ_y	Time-varying probability of recovery from disease (based on rates 0.055–0.192 day^{-1})	Parameter	0.053–0.175
δ_y	Multiplicative lognormal observation error	Parameter	Random
$Z_{\Theta,y}$	Sample size for fishery-independent age composition	Parameter	20 or 200
$Z_{A,y}$	Sample size for antibody prevalence survey	Parameter	20 or 200
EM specific			
$\hat{SB}_y$	Predicted spawning biomass that is available to the survey	Derived quantity	Eq. 19
$\hat{B}_y$	Predicted surveyed biomass	Derived quantity	Eq. 21
$\hat{\Theta}_{y,a}$	Predicted age-composition in fishery-independent sample in year y	Derived quantity	Eq. 22
$\hat{A}_{+,y,a}$	Predicted numbers of fish with antibodies at age a in year y	Derived quantity	Eq. 24
$\hat{A}_{-,y,a}$	Predicted numbers of fish without antibodies at age a in year y	Derived quantity	Eq. 25
$\hat{A}_{y,p}$	Predicted proportions of combined positive and negative antibodies across ages	Derived quantity	Eq. 26
$E_{y,a}$	Proportion of age a fish that are susceptible or exposed in year y	Derived quantity	Eq. 29
$\overline{N}$	Initial average numbers across ages	Parameter	Est
$\tilde{\Omega}_a^*$	Initial proportions of immune fish in each age	Parameter	0 (fixed)
ω_y	Infection rate of susceptible fish in year y	Parameter	$\sum_{d=1}^{n_{day}}(\beta_{I,d,a}+\beta_{C,d,a})v_a S_{y,d,a}/v_a S_{y,d=1,a}$ (Est)
μ	Annual recovery rate of infected fish	Parameter	$\gamma/(\alpha+\gamma) = 0.70$ (Est)
κ	Scaling coefficient for the effect of annual infection prevalence	Parameter	Est
$k_{y,a}$	Observed number of seropositive fish in samples of age a and in year y	Data	—
$\hat{p}_{y,a}$	Predicted proportion of seropositive fish in age a and year y	Derived quantity	$\hat{A}^+_{y,a}/\hat{A}^+_{y,a}+\hat{A}^-_{y,a}$
$N^{sero}_{y,a}$	Total number of fish sampled for seroprevalence in age a and year y	Data	—

Note: Values are grouped by their usage in both the operating and estimation models, in only the operating model, or only the estimation model. Default values (where specified by used) and deriving equations are provided. Estimation model parameters are indicated as fixed or estimated (Est). Default values (or equations in a couple instances) provided for estimation model parameters are the true values specified by the user or computed from variables in the operating model.

Table A2. Equations for the plus-age group in the operating and estimation models. Table 1 provides variable definitions.

Equation	Description
Operating model	
$T_{y+1,d=1,n_a} = \exp(-M_d \cdot 365)\left[(1-\nu_{n_a-1})T_{y,d=1,n_a-1} + (1-\nu_{n_a})T_{y,d=1,n_a}\right] + \exp(-M_d \cdot \Delta t)\left(\nu_{n_a-1}T_{y,n_{\text{day}},n_a-1} + \nu_{n_a}T_{y,n_{\text{day}},n_a}\right)$	Total susceptible numbers in the plus-group at the start of the next season
$S_{y+1,d=1,n_a} = \nu_{n_a}T_{y+1,d=1,n_a}$	Reservoir susceptible numbers in the plus-group at the start of the next season
$I_{y+1,d=1,n_a} = 0$	Infected numbers in the plus-group at the start of the next season
$C_{y+1,d=1,n_a} = \exp(-M_d \cdot \Delta t)\left(C_{y,n_{\text{day}},n_a-1} + C_{y,n_{\text{day}},n_a}\right)$	Carrier numbers in the plus-group at the start of the next season
Estimation model	
$N_{y+1,n_a} = \exp(-M_y)\left\{\left[1-\nu_{n_a-1}E_{y,n_a-1}\omega_y(1-\mu)\right]N_{y,n_a-1} + \left[1-\nu_{n_a}E_{y,n_a}\omega_y(1-\mu)\right]N_{y,n_a}\right\}$	Numbers in plus-group
$\widetilde{\Omega}_{y+1,n_a} = \dfrac{\exp(-M_y)\left(\widetilde{\Omega}_{y,n_a-1} + \nu_{n_a-1}E_{y,n_a-1}\omega_y\mu\right)N_{y,n_a-1} + \exp(-M_y)\left(\widetilde{\Omega}_{y,n_a} + \nu_{n_a}E_{y,n_a}\omega_y\mu\right)N_{y,n_a}}{N_{y+1,n_a}}$	Proportion of immune in plus-group