Using Cure Models for Analyzing the Influence of Pathogens on Salmon Survival

Transactions of the American Fisheries Society 143:387–398, 2014

ISSN: 0002-8487 print / 1548-8659 online
DOI: 10.1080/00028487.2013.862183

ARTICLE

Using Cure Models for Analyzing the Influence of Pathogens on Salmon Survival

R. Adam Ray
Department of Fisheries and Wildlife, Oregon State University, 104 Nash Hall, 2820 Southwest Campus Way, Corvallis, Oregon 97331, USA

Russell W. Perry
U.S. Geological Survey, Western Fisheries Research Center, Columbia River Research Laboratory, 5501A Cook-Underwood Road, Cook, Washington 98605, USA

Nicholas A. Som
U.S. Fish and Wildlife Service, Arcata Fish and Wildlife Office, 1655 Heindon Road, Arcata, California 95521, USA

Jerri L. Bartholomew*
Department of Microbiology, Oregon State University, 220 Nash Hall, 2820 Southwest Campus Way, Corvallis, Oregon 97331, USA

Abstract

Parasites and pathogens influence the size and stability of wildlife populations, yet many population models ignore the population-level effects of pathogens. Standard survival analysis methods (e.g., accelerated failure time models) are used to assess how survival rates are influenced by disease. However, they assume that each individual is equally susceptible and will eventually experience the event of interest; this assumption is not typically satisfied with regard to pathogens of wildlife populations. In contrast, mixture cure models, which comprise logistic regression and survival analysis components, allow for different covariates to be entered into each part of the model and provide better predictions of survival when a fraction of the population is expected to survive a disease outbreak. We fitted mixture cure models to the host–pathogen dynamics of Chinook Salmon *Oncorhynchus tshawytscha* and Coho Salmon *O. kisutch* and the myxozoan parasite *Ceratomyxa shasta*. Total parasite concentration, water temperature, and discharge were used as covariates to predict the observed parasite-induced mortality in juvenile salmonids collected as part of a long-term monitoring program in the Klamath River, California. The mixture cure models predicted the observed total mortality well, but some of the variability in observed mortality rates was not captured by the models. Parasite concentration and water temperature were positively associated with total mortality and the mortality rate of both Chinook Salmon and Coho Salmon. Discharge was positively associated with total mortality for both species but only affected the mortality rate for Coho Salmon. The mixture cure models provide insights into how daily survival rates change over time in Chinook Salmon and Coho Salmon after they become infected with *C. shasta*.

Parasites and pathogens influence the size and structure of host populations (Anderson and May 1978; May and Anderson 1978; Dobson 1988; Hudson et al. 1998), and wildlife populations that have been reduced to low levels may be particularly susceptible to the effects of disease. Dramatic population declines have occurred across a range of species (e.g., corals, amphibians, birds, and mammals) as a result of more frequent and severe disease outbreaks (Harvell et al. 2002). Scientists

*Corresponding author: bartholj@science.oregonstate.edu
Received August 5, 2013; accepted October 26, 2013

use population models (e.g., population viability analysis) to synthesize biological and environmental data in order to predict the long-term stability of a population (Boyce 1992). However, many of these models ignore the population-level effects of pathogens and parasites (Haydon et al. 2002). Incorporation of host–pathogen dynamics into these models can provide critical insights into a population's viability, particularly for species with high cultural, economic, and conservation value (e.g., salmonids).

Predicting a population's long-term viability requires accurate estimates of survival (Gilroy et al. 2012). Standard survival analysis methods (e.g., the Cox proportional hazard model and the accelerated failure time model) assume that each individual in the population will eventually experience the event of interest (Corbière and Joly 2007). Although infections by some wildlife pathogens can result in 100% mortality, in most disease outbreaks some fraction of the population will survive. For example, *Flavobacterium columnare* outbreaks can result in up to 100% mortality of cultured salmonids and catfish, and other pathogens (e.g., viral hemorrhagic septicemia virus and *Myxobolus cerebralis*) have similar wide-ranging effects on both cultured and natural populations (Holt et al. 1975; Thompson et al. 1999; Thomas-Jinu and Goodwin 2004; Kim and Faisal 2010). A different analytical approach is required to quantify factors that may affect the surviving fraction of a population. Cure models, which constitute an extension of survival analysis methods, can be used to investigate heterogeneity between individuals by analyzing the population as two distinct groups: individuals that succumb to disease and individuals that are long-term survivors (Othus et al. 2012).

An example of a heterogeneous population response to a pathogen occurs in populations of juvenile salmon that are infected by the myxozoan parasite *Ceratomyxa shasta*. This parasite is enzootic in major river systems throughout the Pacific Northwest (Margolis and Evelyn 1975; Ratliff 1981; Ching and Munday 1984; Bartholomew 1998), and it is well studied in the Klamath River, California. Data analyzed from a long-term pathogen monitoring study identified a link between *C. shasta* infection and lower abundances of returning adult salmon (Fujiwara et al. 2011). The monitoring program also provided evidence of individual heterogeneity in response to the parasite, as *C. shasta*-induced mortality ranged from 0% to 98% in juvenile Chinook Salmon *Oncorhynchus tshawytscha* and Coho Salmon *O. kisutch* (Hallett et al. 2012; Ray et al. 2012). *Ceratomyxa shasta*-induced mortality in juvenile salmon has three characteristic traits that are difficult to capture by use of traditional survival analysis methods: (1) a delayed onset of mortality after exposure to *C. shasta*; (2) a period of high mortality rate, during which most of the susceptible fish die; and (3) a plateau in the survival curve, where no additional mortality occurs (Figure 1).

The goals of the present study were twofold: we sought to (1) introduce a novel application of a cure model to pathogens in wildlife populations and (2) develop a mixture cure model for predicting *C. shasta*-induced mortality of Chinook Salmon and Coho Salmon in the Klamath River. Our application of the mixture cure model will help us to understand the complex interactions among host, parasite, and environment. Ultimately, these insights will assist in guiding management and conservation actions for the two salmon populations.

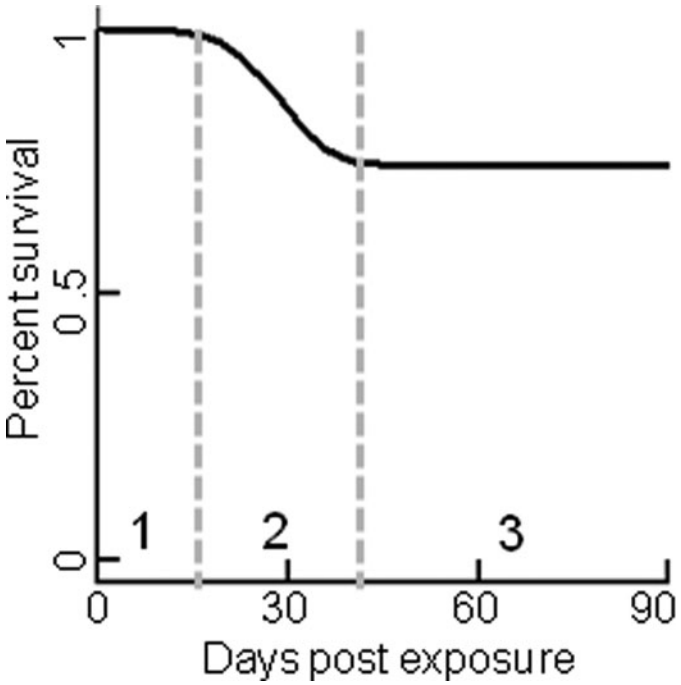

FIGURE 1. Conceptual model depicting the three characteristic traits of *Ceratomyxa shasta*-induced mortality in juvenile salmon: (1) a delayed onset of mortality after exposure to *C. shasta*, (2) a period of high mortality rate, and (3) a plateau in the survival curve, where no additional mortality occurs.

METHODS

Sentinel trials.—The data analyzed here were collected during a long-term project to monitor the spatial and temporal distribution of *C. shasta* and associated parasite-induced mortality in the Klamath River basin. Sentinel trials were conducted by holding juvenile Chinook Salmon and Coho Salmon (age $0+$, obtained from Iron Gate Hatchery, Hornbrook, California) in cages within the Klamath River for 3 d, thus applying *in situ* exposure to varying parasite densities as described by Ray et al. (2012). Although the sentinel trials were conducted at multiple locations, the present analysis focused on data from one site above the confluence with Beaver Creek (see Hallett et al. 2012 for a map), where high parasite densities and high fish mortality were consistently observed (Hallett and Bartholomew 2006; Fujiwara et al. 2011; Hallett et al. 2012). The analyzed data were from sentinel trials conducted with Chinook Salmon ($n = 33$ trials; 1,463 fish exposed) and Coho Salmon ($n = 30$ trials; 1,238 fish exposed) during the summers of 2006–2010 (Table 1).

During the exposure, river temperatures were recorded every 15 min with a HOBO temperature logger (Onset Computer

TABLE 1. Summary of covariate values and observed *Ceratomyxa shasta*-induced mortality for each Chinook Salmon and Coho Salmon sentinel trial used in the mixture cure models (HT = water temperature [°C] during the holding period; ET = water temperature [°C] during the sentinel trial; TI = total Chinook Salmon-specific parasite [genotype I] per liter; TII = total Coho Salmon-specific parasite [genotype II] per liter; Q = discharge [m^3/s] during the sentinel trial). Blank cells for Coho Salmon indicate that no sentinel trial was conducted.

								Chinook Salmon		Coho Salmon	
Trial	Year	Month	HT	ET	TI	TII	Q	Number exposed	Observed mortality (%)	Number exposed	Observed mortality (%)
1	2006	Apr	13	12.2	0.4	0.1	194.9	37	0.0		
2		May	13	18.2	0.0	10.7	168.1	39	0.0	39	5.1
3		Jun	13	20.0	63.2	16.9	118.7	36	19.4	38	2.6
4		Sep	13	20.0	0.2	0.0	34.2	39	0.0		
5	2007	May	13	17.6	4.8	4.8	65.1	39	2.6	30	6.7
6		May	18	17.6	4.8	4.8	65.1	37	27.0	45	86.7
7		Jun	13	20.8	9.7	9.3	53.1	42	2.4	38	2.6
8		Jun	20	20.8	9.7	9.3	53.1	40	40.0	38	81.6
9		Sep	13	20.8	3.2	2.6	31.3	41	7.3	40	2.5
10		Sep	18	20.8	3.2	2.6	31.3	40	2.5	34	35.3
11	2008	May	13	16.2	16.2	6.3	100.2	40	75.0	48	52.1
12		May	16	16.2	16.2	6.3	100.2	41	85.4	41	65.9
13		Jun	13	19.0	42.2	24.8	81.4	77	72.7	75	68.0
14		Jun	15	19.0	42.2	24.8	81.4	75	72.0	79	84.3
15		Jun	18	19.0	42.2	24.8	81.4	76	92.1	70	73.4
16		Jun	21	19.0	42.2	24.8	81.4	86	98.8	74	95.5
17		Sep	13	19.1	7.3	24.5	31.8	41	2.4	35	8.6
18		Sep	18	19.1	7.3	24.5	31.8	39	12.8	19	79.0
19	2009	Apr	13	12.1	37.8	0.0	58.2	36	16.7	42	0.0
20		May	13	14.9	22.0	5.1	62.8	41	73.2	39	12.8
21		May	16	14.9	22.0	5.1	62.8	41	78.1	41	24.4
22		Jun	13	20.8	13.5	3.4	53.8	35	74.3	38	5.3
23		Jun	18	20.8	13.5	3.4	53.8	45	86.7	45	57.8
24		Sep	13	19.9	0.0	0.0	31.1	39	0.0	25	0.0
25		Sep	18	19.9	0.0	0.0	31.1	40	0.0	30	3.3
26	2010	Apr	13	12.6	0.1	0.0	63.8	42	0.0	29	0.0
27		Apr	18	12.6	0.1	0.0	63.8	41	17.1		
28		May	13	13.6	4.0	1.7	70.5	35	0.0	48	0.0
29		May	16	13.6	4.0	1.7	70.5	45	15.6	39	15.4
30		Jun	13	18.2	9.0	1.1	60.4	39	0.0	27	0.0
31		Jun	18	18.2	9.0	1.0	60.4	40	20.0	39	10.3
32		Sep	13	17.0	0.1	0.0	36.0	39	0.0	30	0.0
33		Sep	18	17.0	0.1	0.0	36.0	40	0.0	23	0.0

Corp., Pocasset, Massachusetts) and then were averaged over the 3-d exposure period. To estimate total parasite concentration during exposure, we collected three 1-L samples of river water at the start and end of each sentinel trial; the samples were subsequently filtered and assayed by using a *C. shasta*-specific quantitative PCR technique as described by Hallett et al. (2012). The proportion of parasite genotypes that were specific to Chinook Salmon (type I [TI]) and Coho Salmon (type II [TII]) were also quantified from these samples, as they can influence the severity of infection (Atkinson and Bartholomew 2010a, 2010b; Hallett et al. 2012). We estimated discharge (m^3/s) during the sentinel trials by subtracting tributary discharge (Scott River; U.S. Geological Survey [USGS] gaging station 11519500) from the discharge in the main-stem Klamath River (USGS gaging station 11520500).

After sentinel exposures, fish were held at two temperatures: 13°C (the ambient laboratory water temperature) and an elevated temperature that best represented the in-river conditions during the sentinel trial (15–21°C). The exception was that for trials conducted in 2006, only the ambient temperature was available. Fish were observed for signs of *C. shasta*-induced mortality during a holding period of up to 90 d, and time to mortality

(d) was recorded for each fish. We visually examined each fish for the myxospore stage of *C. shasta* via the methods described by the American Fisheries Society Fish Health Section (AFS-FHS 2012). When myxospores were not visually identified, a section of intestine was assayed by using a *C. shasta*-specific PCR to determine whether fish were infected (Palenzuela and Bartholomew 2002). Fish that survived to the end of the observation period or died from causes other than *C. shasta* were right-censored in the analysis. A right-censored individual is one whose exact time to the event of interest (in this case, *C. shasta*-induced mortality) is unknown, either due to surviving past the observation period or due to being removed from the study by an event other than the event of interest (Collett 2003).

Cure model.—For this analysis, we selected a mixture cure model because it directly models the survival of two distinct groups: individuals that experience the event of interest (susceptible individuals) and individuals that will never experience the event (unsusceptible or "cured" individuals; Othus et al. 2012). In our application, the event of interest is death caused by *C. shasta*. Susceptible individuals are those that die due to *C. shasta*, and cured individuals are those that survive the 90-d holding period. We assumed that salmon could be "cured" either by failing to be infected during the exposure period or by recovering from infection.

A mixture cure model is a survival distribution function that combines (1) a logistic model for the probability of death and (2) a standard survival model for the time to death of susceptible individuals (Othus et al. 2012),

$$S(t|\boldsymbol{x}, \boldsymbol{z}) = [1 - \pi(\boldsymbol{z})] + \pi(\boldsymbol{z})S(t|U = 1, \boldsymbol{x}), \qquad (1)$$

where $S(t|\boldsymbol{x}, \boldsymbol{z})$ is the survival distribution function for time t given covariate vectors $\boldsymbol{x}$ and $\boldsymbol{z}$; $S(t|U = 1, \boldsymbol{x})$ is the survival distribution function for time t conditional on individuals that died due to *C. shasta*; U is an indicator function of *C. shasta*-induced mortality ($U = 1$ for fish that died due to *C. shasta* and 0 otherwise); and $\pi(\boldsymbol{z})$ is the probability of death due to *C. shasta* (Corbière and Joly 2007). We used the Weibull distribution for $S(t|U = 1, \boldsymbol{x})$ because it is flexible and contains a number of other distributions as special cases (e.g., the exponential distribution).

Covariates can be included in the mixture cure model, where $\boldsymbol{z}$ is the covariate vector for the proportion of individuals that die and $\boldsymbol{x}$ is the covariate vector for the timing and rate of mortality among susceptible individuals. The same covariates can be included in both $\boldsymbol{x}$ and $\boldsymbol{z}$, but this is not required. For the logistic portion of the model, the logit of $\pi(\boldsymbol{z})$ is expressed as a linear function of the covariates. For the survival distribution part of the model, $S(t|U = 1, \boldsymbol{x})$ is implemented as an accelerated failure time model in which $\log(t)$ is modeled as a linear function of the covariates (Peng et al. 1998).

The covariates of interest for our models included the total amount of host-specific parasite per liter (TI for Chinook Salmon; TII for Coho Salmon), the water temperature during the holding period at the laboratory (HT), the water temperature during the 3-d exposure period (ET), the average discharge during the 3-d sentinel trial (Q), and the interactions TI $\times$ HT and TI $\times$ Q for Chinook Salmon and TII $\times$ HT and TII $\times$ Q for Coho Salmon. For each of the two salmon species, we developed separate global models that included all of the covariates and interaction terms in both model components (Table 2). To improve model convergence, we centered the covariates by subtracting the mean from each observation.

Covariates and interaction terms were selected based on their hypothesized effects on *C. shasta*-induced mortality. The concentration of species-specific parasite (TI for Chinook Salmon; TII for Coho Salmon) was calculated by multiplying the total concentration of parasite DNA by the proportion of each genotype present in the water (Hallett et al. 2012). Water temperature is associated with higher total mortality and faster mortality rates in both Chinook Salmon and Coho Salmon, so we included two covariates for water temperature: (1) ET (temperature during the exposure period) and (2) HT (temperature during the holding period; Udey et al. 1975; Hallett et al. 2012; Ray et al. 2012). We included Q as a proxy for velocity because Ray and Bartholomew (2013) identified an inverse relationship between water velocity and parasite attachment to the gills. The interactions TI $\times$ HT (for Chinook Salmon) and TII $\times$ HT (for Coho Salmon) account for the known compounding effects of parasite concentration and water temperature on the mortality rate and total mortality (Ray et al. 2012), while the interactions TI $\times$ Q and TII $\times$ Q act as proxies for the total exposure dose (Hallett et al. 2012).

Parameters for the candidate models were estimated via maximum likelihood with the gfcure package in R (Zhang and Peng 2007; R Development Core Team 2011). The weight of evidence for each model was then assessed by using Akaike's information criterion (AIC) to identify the most parsimonious model (Akaike 1973; Burnham and Anderson 2002). Candidate models were constructed by first separately removing the ET covariate from the logistic component and then from the survival component, as the exposure period constituted only a small fraction of the entire monitoring period. Selecting the model with the lowest AIC value, we then individually removed each interaction term. After removing a covariate, we would select the model with the lowest AIC value and repeat this process until the model with the lowest AIC score was selected as the final model.

We present the results of the final models in two ways. First, the estimated Kaplan–Meier (KM) survivor function is plotted for each sentinel trial (Collett 2003), and the predicted curves from the final models are compared. To assess the accuracy of the logistic regression component, we calculated a Brier score, which ranges from 0.00 (perfect fit) to 0.25 (poor fit); we assumed that a score less than 0.125 indicated a good fit (Steyerberg et al. 2001). To assess the fit of the survival model, we plotted 95% confidence intervals of the predicted cure model curves. Second, we evaluate the relative influence of the HT and Q covariates on parasite-induced mortality by plotting the minimum, mean, and maximum observed values of each

TABLE 2. Model selection results for Chinook Salmon and Coho Salmon mixture cure models based on the Weibull distribution. All covariates are shown for the global model, with other models showing terms that were removed from the global model. Blank cells indicate that no covariates were removed ($\pi[z]$ = logistic model; $S[t]$ = survival model; k = number of estimated parameters; MLL = maximized log-likelihood; AIC = Akaike's information criterion; ΔAIC = AIC difference). The final model for each species is presented in bold type.

Number	Component	Model	k	MLL	AIC	ΔAIC
		Chinook Salmon				
Global	$\pi(z)$	HT + ET + TI + Q + (TI × HT) + (TI × Q)	15	−519.93	1,069.87	1.53
	$S(t)$	HT + ET + TI + Q + (TI × HT) + (TI × Q)				
1	$\pi(z)$		14	−561.66	1,151.32	82.98
	$S(t)$	ET removed				
2	$\pi(z)$	ET removed	14	−520.54	1,069.08	0.74
	$S(t)$					
3	$\pi(z)$	ET, (TI × HT) removed	13	−533.40	1,092.81	24.47
	$S(t)$					
4	$\pi(z)$	ET, (TI × Q) removed	13	−525.68	1,077.36	9.02
	$S(t)$					
5	$\pi(z)$	ET removed	13	−616.85	1,259.70	191.36
	$S(t)$	(TI × Q) removed				
Final	$\boldsymbol{\pi(z)}$	**ET removed**	**13**	**−521.17**	**1,068.34**	**0.00**
	$\boldsymbol{S(t)}$	**(TI × HT) removed**				
6	$\pi(z)$	ET, (TI × HT) removed	12	−533.96	1,091.92	23.58
	$S(t)$	(TI × HT) removed				
7	$\pi(z)$	ET, (TI × Q) removed	12	−526.30	1,076.60	8.26
	$S(t)$	(TI × HT) removed				
8	$\pi(z)$	ET removed	12	−619.32	1,262.65	194.31
	$S(t)$	(TI × HT), (TI × Q) removed				
		Coho Salmon				
Global	$\pi(z)$	HT + ET + TII + Q + (TII × HT) + (TII × Q)	15	−662.51	1,355.02	3.27
	$S(t)$	HT + ET + TII + Q + (TII × HT) + (TII × Q)				
1	$\pi(z)$		14	−664.79	1,357.59	6.36
	$S(t)$	ET removed				
2	$\pi(z)$	ET removed	14	−663.48	1,354.96	3.73
	$S(t)$					
3	$\pi(z)$	ET, (TII × Q) removed	13	−667.01	1,360.02	8.79
	$S(t)$					
4	$\pi(z)$	ET, (TII × HT) removed	13	−665.51	1,353.01	1.78
	$S(t)$					
5	$\pi(z)$	ET removed	13	−674.55	1,357.10	5.87
	$S(t)$	(TII × HT) removed				
6	$\pi(z)$	ET removed	13	−663.61	1,353.21	1.98
	$S(t)$	(TII × Q) removed				
7	$\pi(z)$	ET, (TII × HT), (TII × Q) removed	12	−667.06	1,358.12	6.89
	$S(t)$					
8	$\pi(z)$	ET, (TII × HT) removed	12	−665.80	1,355.60	4.37
	$S(t)$	(TII × HT) removed				
Final	$\boldsymbol{\pi(z)}$	**ET, (TII × HT) removed**	**12**	**−663**	**1,351**	**0**
	$\boldsymbol{S(t)}$	**(TII × Q) removed**				
9	$\pi(z)$	ET, (TII × HT), (TII × Q) removed	11	−667.06	1,356.13	4.90
	$S(t)$	(TII × Q) removed				
10	$\pi(z)$	ET, (TII × HT) removed	11	−666.60	1,355.19	3.96
	$S(t)$	(TII × Q), (TII × HT) removed				

covariate against the minimum, mean, and maximum observed parasite concentrations (TI for Chinook Salmon; TII for Coho Salmon).

RESULTS

Chinook Salmon

We found that the best-fitting model based on AIC excluded ET from the logistic component and excluded TI × HT from the survival component (AIC = 1,068.34; Table 2). Removal of the ET covariate from the logistic component (model 2; AIC = 1,069.08) resulted in an AIC value lower than that obtained by the removal of ET from the survival component (model 1; AIC = 1,151.32) and lower than that of the global model (AIC = 1,069.87). Removing the TI × HT term from the survival component (final model) produced the lowest AIC value. Individual removal of the remaining interaction terms (models 6–8) resulted in AIC scores that were higher than the AIC of the final model.

In general, the Chinook Salmon mixture cure model was able to capture all three of the mortality characteristics (the delayed onset of mortality, the period of high mortality, and the plateau in which no further mortality occurs) for a majority of the sentinel trials (Figure 2). The model correctly predicted the onset of mortality for almost all of the sentinel trials in which mortality was observed. The logistic component of the model accurately (Brier score < 0.125) reproduced 84.8% (28/33) of the estimated total mortalities. Three trials (i.e., trials 11, 20, and 23) had Brier scores between 0.125 and 0.25, suggesting an adequate fit between predicted and observed data, but the remaining two trials (i.e., trials 3 and 22) had Brier scores greater than 0.25, indicating a poor model fit. The survival component replicated 69.7% (23/33) of the observed mortality rates (i.e., slopes). In general, the mixture cure model best reproduced the slopes of the KM curves when mortality exceeded 50% (e.g., trials 12–16). The model always predicted mortality even when none was observed (e.g., trials 1, 2, and 30). Our model was able to reproduce at least one of the three mortality characteristics for all sentinel trials except trial 3, for which none of the observed patterns was captured.

In the final model, all of the logistic regression covariates were positively associated with the probability of mortality due to *C. shasta*, and all of the survival analysis covariates were negatively associated with the survival rate (Table 3). Overall, both the total mortality and the rate of mortality in Chinook Salmon increased as the TI value increased (from left to right, Figure 3); however, this response differed with the interacting covariate: either HT (Figure 3a–c) or Q (Figure 3d–f). Although the total mortalities were similar between HT and Q for each value of TI, increasing the value of Q resulted in higher total mortality than increasing the value of HT. The mortality rate was more influenced by increasing the value of HT than by increasing Q.

TABLE 3. Parameter coefficients from final mixture cure models (see Table 2) for both Chinook Salmon and Coho Salmon.

Component	Covariate	Coefficient	SE
	Chinook Salmon		
Logistic	Intercept	−0.280	0.101
	HT	0.389	0.045
	TI	0.123	0.010
	Q	0.022	0.005
	TI × HT	0.019	0.004
	TI × Q	0.001	0.0003
Survival	Intercept	3.424	0.013
	HT	−0.084	0.003
	ET	−0.037	0.005
	TI	−0.007	0.001
	Q	−0.003	0.001
	TI × Q	−0.001	0.0001
	Log (scale)	−1.561	0.030
	Coho Salmon		
Logistic	Intercept	−0.555	0.091
	HT	0.524	0.045
	TII	0.115	0.009
	Q	0.012	0.003
	TII × Q	−0.001	0.0004
Survival	Intercept	3.907	0.034
	HT	−0.134	0.018
	ET	0.038	0.153
	TII	−0.020	0.003
	Q	−0.006	0.001
	TII × HT	0.003	0.001
	Log (scale)	−1.136	0.034

Coho Salmon

We found that the best-fitting Coho Salmon model based on AIC excluded ET and TII × HT from the logistic component and excluded TII × Q from the survival component (AIC = 1,351.75; Table 2). Removal of the ET covariate from the logistic regression component (model 2; AIC = 1,354.96) resulted in an AIC score lower than that observed for removal of ET from the survival component (model 1; AIC = 1,357.591) and lower than that of the global model (AIC = 1,355.02). Removing the TI × HT term from the logistic component and removing TI × Q from the survival component resulted in the lowest AIC value (final model). Removal of the two remaining interaction terms increased the AIC score (models 9 and 10).

Although Coho Salmon respond differently than Chinook Salmon to *C. shasta* infection, the Coho Salmon mixture cure model was able to capture all three mortality characteristics for a majority of these sentinel trials (Figure 4). The model predicted the onset of mortality for a majority of sentinel trials except when no mortality was observed (e.g., trial 28) or when the mortality onset occurred much later than average (e.g., trial 25).

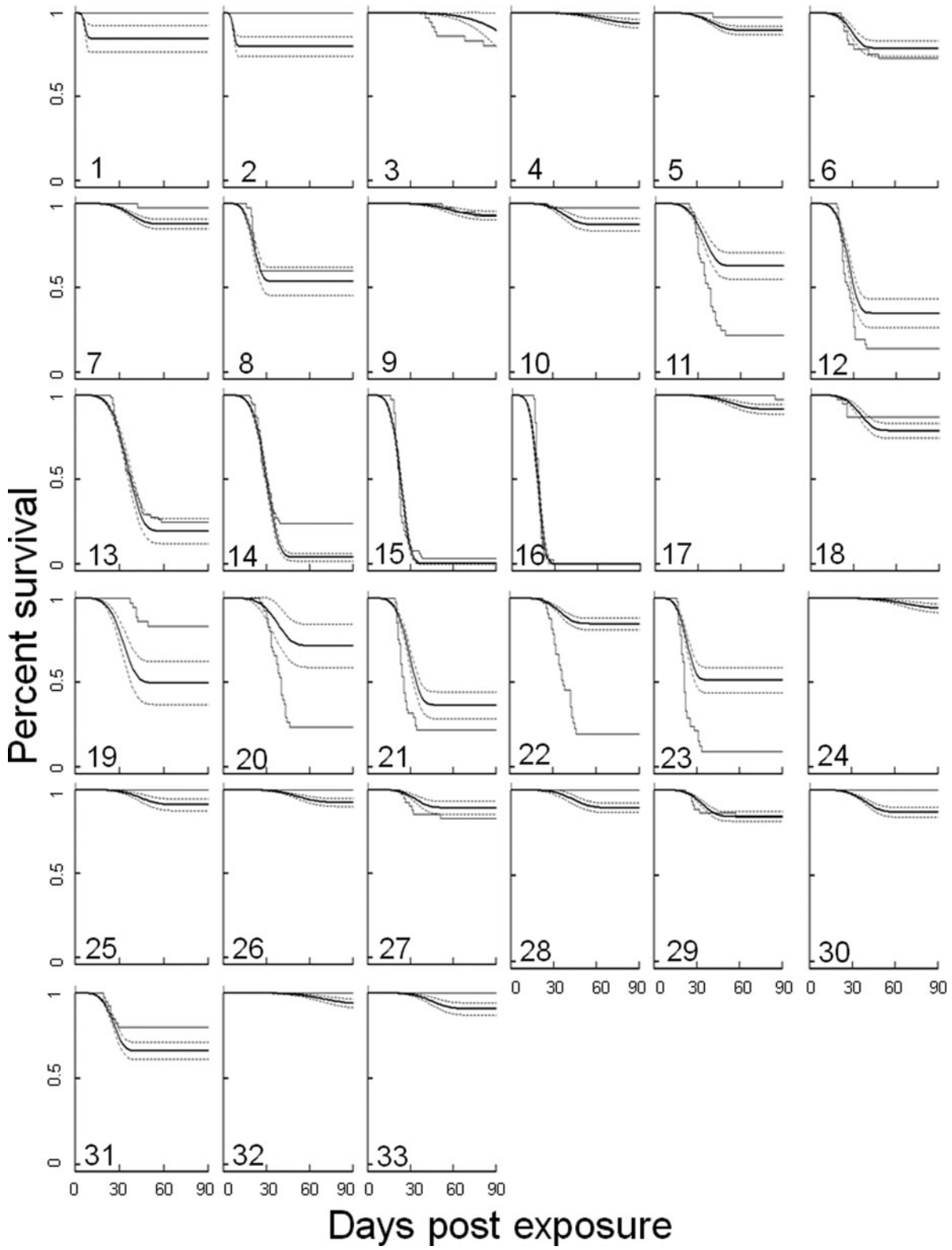

FIGURE 2. Estimated Kaplan–Meier (thin lines) and mixture cure model (bold lines) survival curves (with 95% confidence intervals, dashed lines) for Chinook Salmon sentinel trials conducted in the Klamath River, California. Numbers in the lower left corner of each panel correspond to the trial numbers described in Table 1.

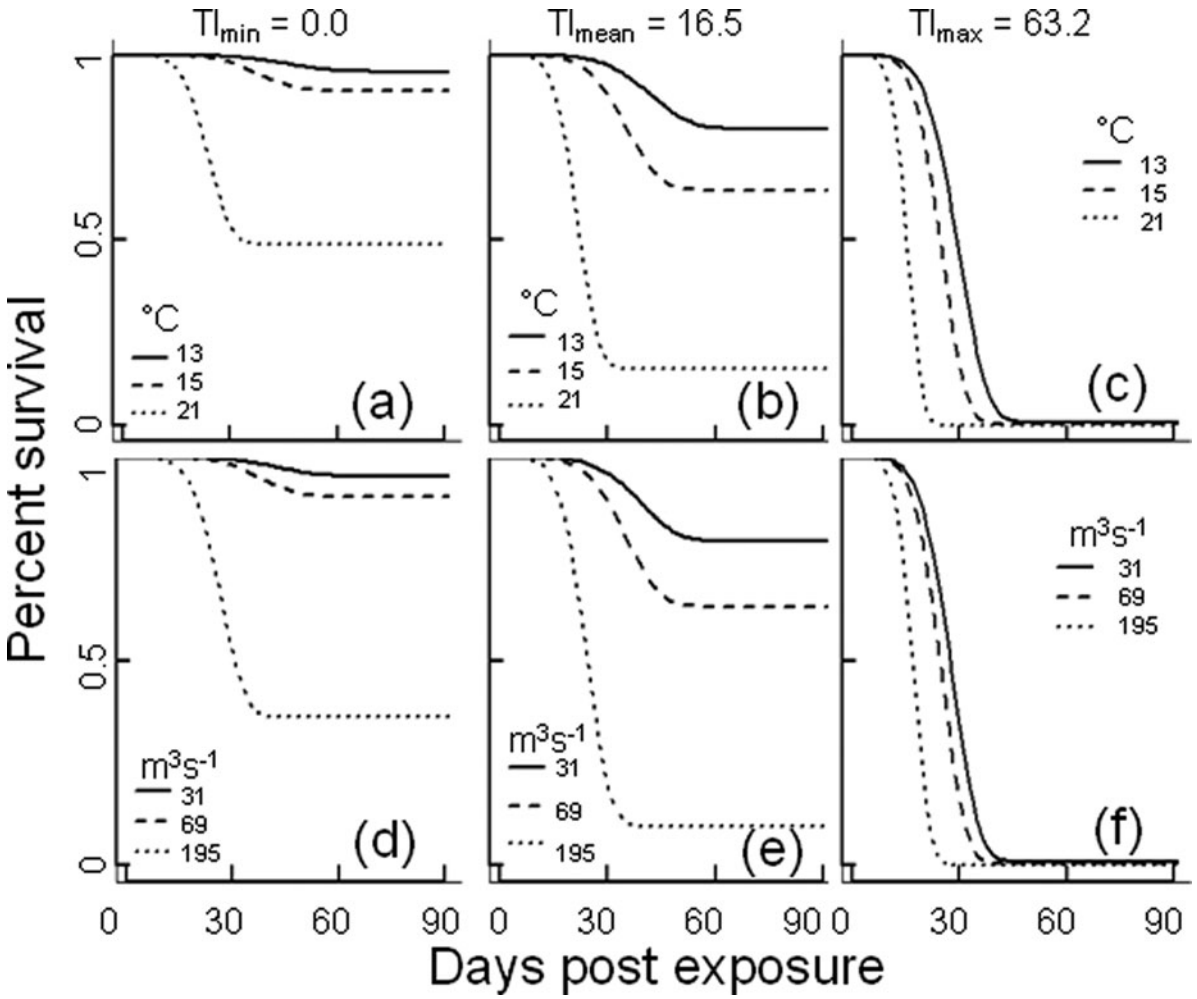

FIGURE 3. Response of predicted survival probability from the Chinook Salmon mixture cure model at minimum (a and d), mean (b and e), and maximum (c and f) concentrations of Chinook Salmon-specific *Ceratomyxa shasta* (genotype I [TI] per liter). The top row (a–c) represents the interacting effect of holding temperature (HT) and TI on the predicted survival probability. The bottom row (d–f) represents the interacting effect of discharge (Q) and TI on the predicted survival probability. The lines within each panel correspond to the minimum (solid line), mean (dashed line), and maximum (dotted line) observed values of HT and Q.

The logistic component of the model accurately (Brier score < 0.125) reproduced the observed total mortality for 93.3% (28/30) of the sentinel trials. Unlike the Chinook Salmon model, all Brier scores for the Coho Salmon model were less than 0.25, but there were two trials (i.e., trials 11 and 17) that exceeded our threshold score of 0.125. The survival component of the model was able to replicate 56.7% (17/30) of the estimated KM curves, especially when total mortality was greater than 50% (e.g., trials 12–16). The model was able to predict zero mortality and low mortality in some cases (e.g., trials 5, 19, and 32) but was inconsistent (predicting either faster or slower mortality) for other trials. The Coho Salmon mixture cure model was able to predict the overall mortality for almost all of the trials (93.3%), but it did not perform as well as the Chinook Salmon model at replicating the mortality rate in the observed trials. However, the Coho Salmon model was able to reproduce at least one of the three mortality characteristics for all observed sentinel trials.

Similar to the Chinook Salmon model, the HT, TII, and Q covariates in the logistic component of the Coho Salmon model were positively associated with the probability of total *C. shasta*-induced mortality; however, the term TII $\times$ Q was associated with a decreased probability of total mortality (Table 3). In the survival analysis component of the model, all of the covariates (except ET and TII $\times$ HT) were negatively associated with survival rate. Overall, increases in TII produced increases in both total mortality and the mortality rate; as with Chinook Salmon, the pattern of this response for Coho Salmon also differed between the HT (Figure 5a–c) and Q (Figure 5d–f) covariates. Total mortality and the rate of mortality were greater at minimum values of Q than at minimum values of HT across all TII values. However, increasing the value of HT resulted in greater

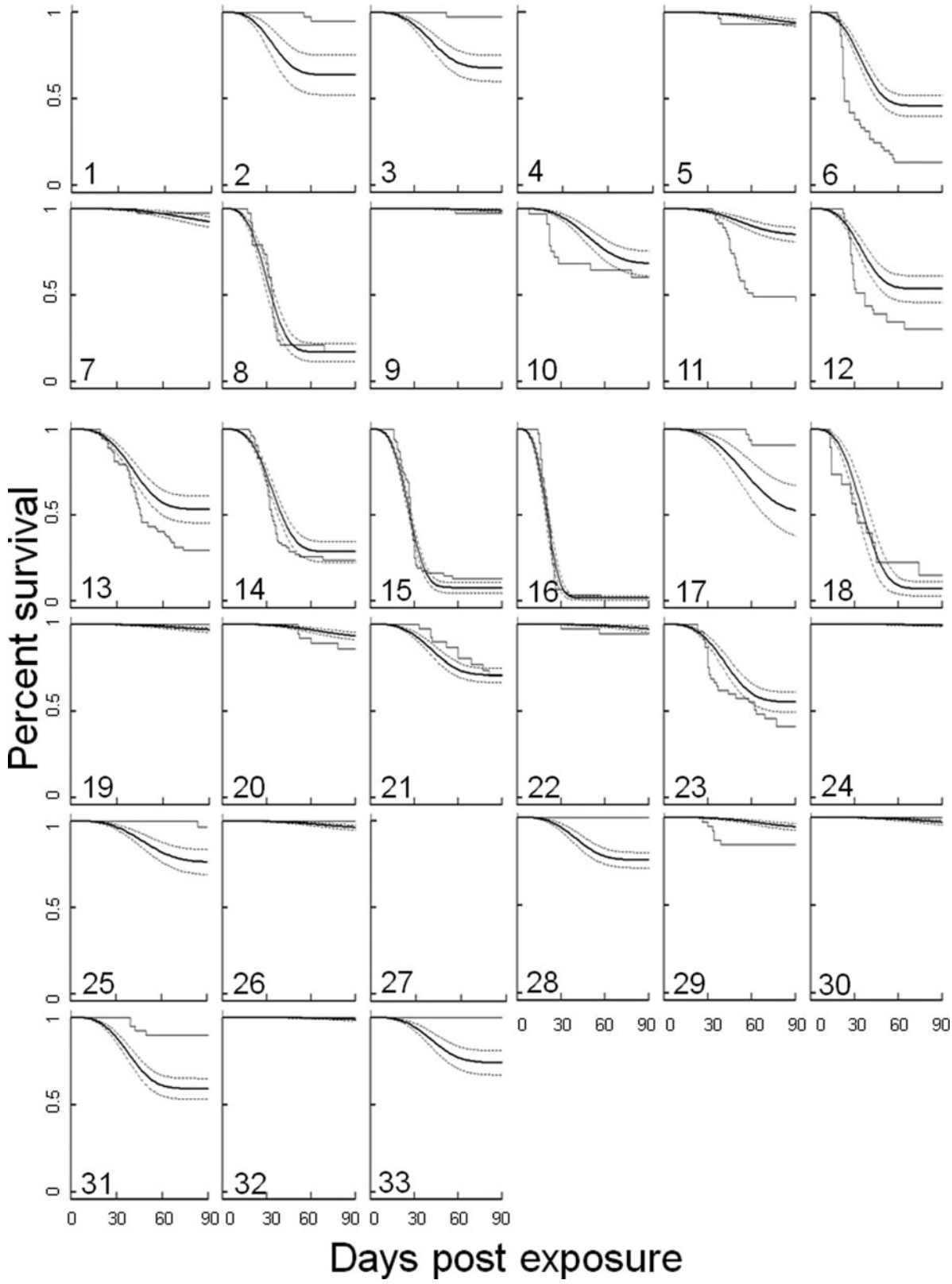

FIGURE 4. Estimated Kaplan–Meier (thin lines) and mixture cure model (bold lines) survival curves (with 95% confidence intervals, dashed lines) for Coho Salmon sentinel trials conducted in the Klamath River. Numbers in the lower left corner of each panel correspond to the trial numbers described in Table 1. Trials 1, 4, and 27 were not conducted with Coho Salmon.

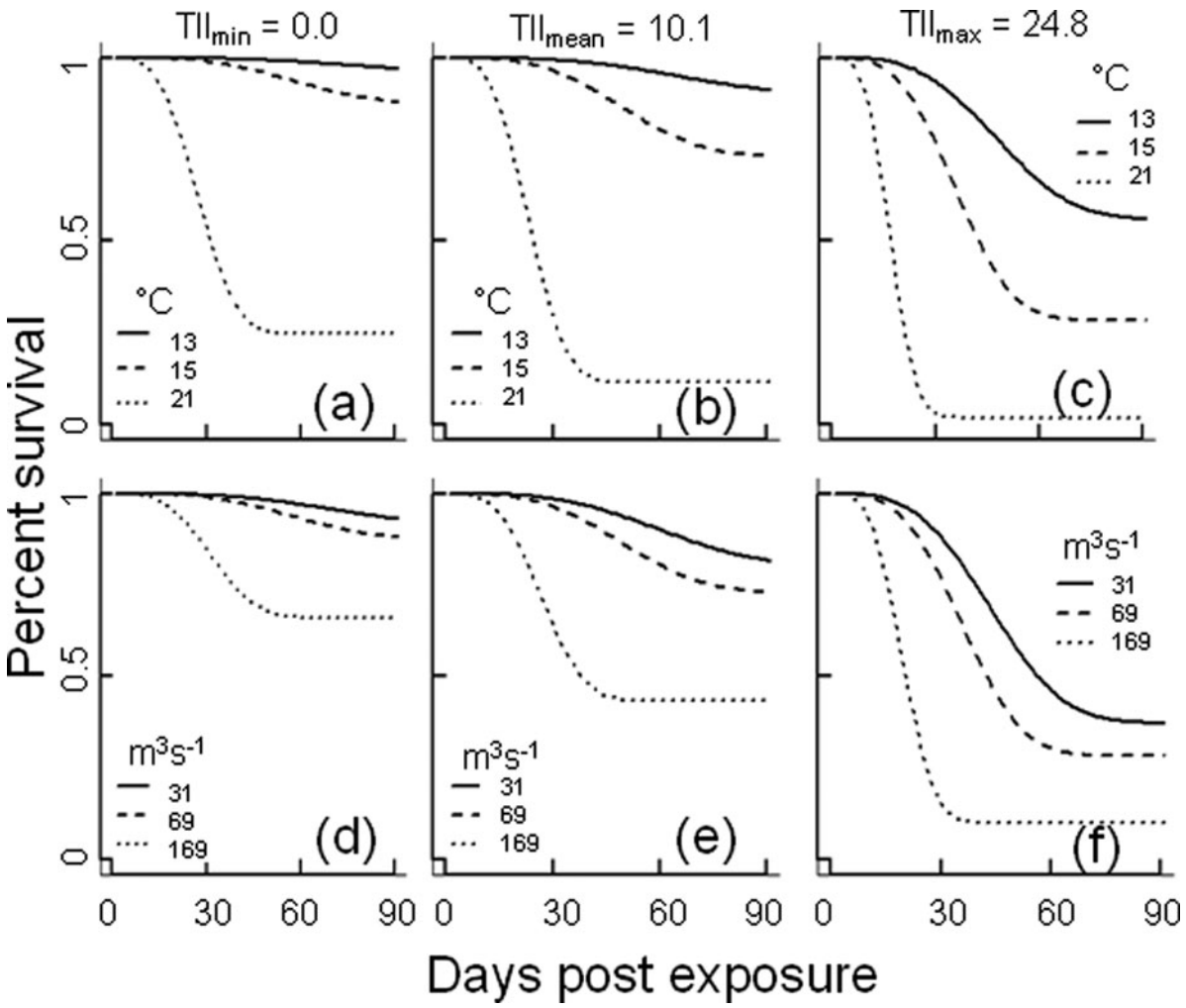

FIGURE 5. Response of predicted survival probability from the Coho Salmon mixture cure model at minimum (a and d), mean (b and e), and maximum (c and f) concentrations of Coho Salmon-specific *Ceratomyxa shasta* (genotype II [TII] per liter). The top row (a–c) represents the interacting effect of holding temperature (HT) and TII on the predicted survival probability. The bottom row (d–f) represents the interacting effect of discharge (Q) and TII on the predicted survival probability. The lines within each panel correspond to minimum (solid line), mean (dashed line), and maximum (dotted line) observed values of HT and Q.

differences in both total mortality and the mortality rate than increasing the value of Q.

DISCUSSION

Cure models provide an analytical tool for modeling and predicting survival rates in wildlife populations, especially those in which abundance is constrained by pathogens. Cure models can be analyzed by parametric or semiparametric methods, thereby allowing for use in a wide array of host–pathogen systems with varying amounts of empirical data. In this paper, we presented a parametric version of the model, as it allows for more precise estimates of the covariates of interest (Collett 2003). Another advantage of using cure models over other survival analysis methods is the ability to incorporate different covariates into the survival and logistic components, further expanding the applications to different disease systems. We used a mixture cure model to analyze the effects of *C. shasta* on the survival rates of Chinook Salmon and Coho Salmon because this model allowed us to divide our populations into two distinct groups—those that succumb to infection (susceptible individuals) and those that survive infection ("cured" individuals).

In this host–pathogen system, the "cured" proportion of the salmon host can arise in two different ways. First, not all individuals may become infected with *C. shasta*. When parasite densities are relatively low (<10 parasite/L), parasite-induced mortality is often lower and highly variable (Hallett et al. 2012). Ray and Bartholomew (2013) observed density-independent transmission dynamics between *C. shasta* and its salmon host when parasite densities were low (<10 parasite/L), suggesting that random chance encounters led to infection. However, as parasite densities increased, density-dependent transmission dynamics were observed, indicating that at some parasite concentration threshold almost all salmon will become infected.

Second, salmon may be able to recover from the infection. Evidence for the ability to recover from infection was reported by Ray et al. (2010), who conducted caged exposures in the Klamath River when total *C. shasta* densities exceeded 100 parasites/L. Observed mortality was over 90%, yet no evidence of infection (by either visual inspection or PCR) was detected among the survivors.

In our study, the mixture cure models for Chinook Salmon and Coho Salmon had slightly different structures in terms of the retained covariates, indicating that these salmonid species have biologically different responses to infection. Although several comparable models (ΔAIC < 2) were found for both Chinook Salmon and Coho Salmon, they shared an overall structure similar to that of the final model selected. The logistic component of the final models differed in that TI $\times$ HT was retained in the Chinook Salmon model, but TII $\times$ HT was dropped from the Coho Salmon model. This finding suggests that *C. shasta* TI (Chinook Salmon) could be more virulent than TII (Coho Salmon) and that this virulence is compounded by increased water temperatures, as was hypothesized by Hallett et al. (2012). The survival component of the final model was similarly structured for the two species, except that TI $\times$ Q was retained in the Chinook Salmon model, whereas TII $\times$ Q was removed from (and TII $\times$ HT was retained in) the Coho Salmon model. As TI is the dominant *C. shasta* genotype in this system, the TI $\times$ Q term provides an exposure dose estimate for Chinook Salmon. Retention of the TII $\times$ HT interaction in the survival component for Coho Salmon indicates that rising water temperatures increase the severity and rate of *C. shasta*-induced mortality (Udey et al. 1975; Ray et al. 2012); inclusion of this term in the model suggests that Coho Salmon are more sensitive to warmer temperatures than Chinook Salmon in the Klamath River (Richter and Kolmes 2005; Hallett et al. 2012).

Overall, the mixture cure models captured the observed mortality characteristics for both Chinook Salmon and Coho Salmon; however, the lack of fit among certain trials could be improved by further refining our measurements of environmental variables and by improving our understanding of infection dynamics in the salmon hosts. The estimates of TI for Chinook Salmon and TII for Coho Salmon were based on three 1-L samples collected at the start and end of the exposure period. Continuous collection of water samples during the 3-d exposure period may help to improve the estimates of this important covariate. Although Q was a significant covariate in both models, it is a relatively coarse proxy for water velocity and does not capture variation in velocity among other locations in the river (e.g., eddies, pools, and riffles). Ray and Bartholomew (2013) observed an inverse relationship between velocity and actinospore transmission with a threshold at approximately 0.3 m/s, above which *C. shasta* transmission was greatly reduced. Therefore, a more fine-scale measurement (e.g., near-cage velocity) may improve the fit of these models to the observed data. In addition to more fine-scale measurements of total parasite concentration and velocity, the use of replicate cages would aid in capturing variation in the mortality response, especially during years of low to moderate (<50%) mortality. Atkinson and Bartholomew (2010a) demonstrated a link between the relative proportion of species-specific *C. shasta* genotype and mortality. However, it is not known how co-infection with both genotypes would influence the severity of disease. When TI was the dominant genotype in the water column, Chinook Salmon mortality was generally high (e.g., trials 22 and 23). However, when the proportions of TI and TII were similar (e.g., trials 5 and 7), Chinook Salmon mortality was lower, suggesting that higher proportions of TII may lessen the lethal effects of TI on Chinook Salmon. In addition to multiple genotypes of *C. shasta*, these salmon are also exposed to other pathogens (e.g., *Flavobacterium columnare* and *Aeromonas salmonicida*) and parasites (e.g., *Nanophyetus salmincola*) that could affect overall mortality and the rate of mortality. Refinement of our techniques for measuring environmental factors and the inclusion of interactions among *C. shasta* genotypes, other pathogens, and the salmon hosts may improve the predictive capabilities of these mixture cure models.

As the incidence of epizootics and diseases affecting wildlife populations increases across all taxonomic classes (Harvell et al. 2002), disease ecologists employ a variety of statistical methods and models to understand and quantify host–pathogen dynamics. However, standard survival analysis methods largely ignore the importance of heterogeneity in a population's response to infection. The heterogeneity displayed in wildlife populations makes cure models an attractive alternative to traditional methods of survival analysis. The mixture cure models developed here provide daily survival rates and estimates of population-level parasite-induced mortality that can be incorporated into salmon population and production models. Our mixture cure models allow for a detailed understanding of juvenile salmon survival in the Klamath River, which in turn will allow managers to better account for the effects of disease dynamics on these stocks. Although the present mixture cure models were developed for an aquatic pathogen, they could also be applied to other host–pathogen systems, such as sylvatic plague (*Yersinia pestis*) in prairie dog *Cynomys* spp. colonies, white-nose syndrome (*Geomyces destructans*) in bats, or chytridiomycosis (*Batrachochytrium dendrobatidis*) in amphibians (Cully and Williams 2001; Frick et al. 2010; Muths et al. 2011). Cure models provide a flexible yet powerful analytical tool that can be applied to a wide range of host–pathogen systems, allowing the identification and quantification of biotic and abiotic factors that are significant for both the affected and surviving fractions of the population.

ACKNOWLEDGMENTS

We thank Julie Alexander for providing feedback and support in the development and analysis of the models. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the USGS or the U.S. Fish and Wildlife Service. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. This project was supported by funding from the U.S. Bureau of Reclamation through Cooperative

Agreement R09AC20022, Cooperative Ecosystems Study Unit 3FC810873.

REFERENCES

AFS-FHS (American Fisheries Society, Fish Health Section). 2012. Fish Health Section blue book: suggested procedures for the detection and identification of certain finfish and shellfish pathogens, 2012 edition. American Fisheries Society, Bethesda, Maryland.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pages 267–281 *in* B. N. Petrov and F. Csaki, editors. Second international symposium on information theory. Akadémia Kiadó, Budapest.

Anderson, R. M., and R. M. May. 1978. Regulation and stability of host–parasite population interactions: I. Regulatory processes. Journal of Animal Ecology 47:219–247.

Atkinson, S. D., and J. L. Bartholomew. 2010a. Disparate infection patterns of *Ceratomyxa shasta* (Myxozoa) in Rainbow Trout (*Oncorhynchus mykiss*) and Chinook Salmon (*Oncorhynchus tshawytscha*) correlate with internal transcribed spacer-1 sequence variation in the parasite. International Journal for Parasitology 40:599–604.

Atkinson, S. D., and J. L. Bartholomew. 2010b. Spatial, temporal and host factors structure the *Ceratomyxa shasta* (Myxozoa) population in the Klamath River basin. Infection, Genetics and Evolution 10:1019–1026.

Bartholomew, J. L. 1998. Host resistance to infection by the myxosporean parasite *Ceratomyxa shasta*: a review. Journal of Aquatic Animal Health 10:112–120.

Boyce, M. S. 1992. Population viability analysis. Annual Review of Ecology and Systematics 23:481–506.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer Verlag, New York.

Ching, H. L., and D. R. Munday. 1984. Geographic and seasonal distribution of the infectious stage of *Ceratomyxa shasta* Noble, 1950, a myxozoan salmonid pathogen in the Fraser River system. Canadian Journal of Zoology 62:1075–1080.

Collett, D. 2003. Modelling survival data in medical research. Chapman and Hall/CRC Press, London.

Corbière, F., and P. Joly. 2007. A SAS macro for parametric and semiparametric mixture cure models. Computer Methods and Programs in Biomedicine 85:173–180.

Cully, J. F., Jr., and E. S. Williams. 2001. Interspecific comparisons of sylvatic plague in prairie dogs. Journal of Mammalogy 82:894–905.

Dobson, A. P. 1988. The population biology of parasite-induced changes in host behavior. Quarterly Review of Biology 63:139–165.

Frick, W. F., J. F. Pollock, A. C. Hicks, K. E. Langwig, D. S. Reynolds, G. G. Turner, C. M. Butchkoski, and T. H. Kunz. 2010. An emerging disease causes regional population collapse of a common North American bat species. Science 329:679–682.

Fujiwara, M., M. S. Mohr, A. Greenberg, J. S. Foott, and J. L. Bartholomew. 2011. Effects of ceratomyxosis on population dynamics of Klamath fall-run Chinook Salmon. Transactions of the American Fisheries Society 140:1380–1391.

Gilroy, J. J., T. Virzi, R. L. Boulton, and J. L. Lockwood. 2012. A new approach to the "apparent survival" problem: estimating true survival rates from mark-recapture studies. Ecology 93:1509–1516.

Hallett, S. L., and J. L. Bartholomew. 2006. Application of a real-time PCR assay to detect and quantify the myxozoan parasite *Ceratomyxa shasta* in river water samples. Diseases of Aquatic Organisms 71:109–118.

Hallett, S. L., R. A. Ray, C. N. Hurst, R. A. Holt, G. R. Buckles, S. D. Atkinson, and J. L. Bartholomew. 2012. Density of the waterborne parasite *Ceratomyxa shasta* and its biological effects on salmon. Applied and Environmental Microbiology 78:3724–3731.

Harvell, C. D., C. E. Mitchell, J. R. Ward, S. Altizer, A. P. Dobson, R. S. Ostfeld, and M. D. Samuel. 2002. Climate warming and disease risks for terrestrial and marine biota. Science 296:2158–2162.

Haydon, D. T., M. K. Laurenson, and C. Sillero-Zubiri. 2002. Integrating epidemiology into population viability analysis: managing the risk posed by rabies and canine distemper to the Ethiopian wolf. Conservation Biology 16:1372–1385.

Holt, R. A., J. E. Sanders, J. L. Zinn, J. L. Fryer, and K. S. Pilcher. 1975. Relation of water temperature to *Flexibacter columnaris* infection in steelhead trout (*Salmo gairdneri*), Coho (*Oncorhynchus kisutch*) and Chinook (*O. tshawytscha*) Salmon. Journal of the Fisheries Board of Canada 32:1553–1559.

Hudson, P. J., A. P. Dobson, and D. Newborn. 1998. Prevention of population cycles by parasite removal. Science 282:2256–2258.

Kim, R., and M. Faisal. 2010. Experimental studies confirm the wide host range of the Great Lakes viral haemorrhagic septicaemia virus genotype IVb. Journal of Fish Diseases 33:83–88.

Margolis, L., and T. P. T. Evelyn. 1975. *Ceratomyxa shasta* (Myxosporida) disease in Chum Salmon (*Oncorhynchus keta*) in British Columbia. Journal of the Fisheries Board of Canada 32:1640–1643.

May, R. M., and R. M. Anderson. 1978. Regulation and stability of host–parasite population interactions: II. Destabilizing processes. Journal of Animal Ecology 47:249–267.

Muths, E., R. D. Scherer, and D. S. Pilliod. 2011. Compensatory effects of recruitment and survival when amphibian populations are perturbed by disease. Journal of Applied Ecology 48:873–879.

Othus, M., B. Barlogie, M. L. LeBlanc, and J. J. Crowley. 2012. Cure models as a useful statistical tool for analyzing survival. Clinical Cancer Research 18:3731–3736.

Palenzuela, O., and J. L. Bartholomew. 2002. Molecular tools for the diagnosis of *Ceratomyxa shasta* (Myxozoa). Pages 285–298 *in* C. O. Cunningham, editor. Molecular diagnosis of salmonid diseases. Springer, New York.

Peng, Y., K. B. G. Dear, and J. W. Denham. 1998. A generalized *F* mixture model for cure rate estimation. Statistics in Medicine 17:813–830.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Ratliff, D. E. 1981. *Ceratomyxa shasta*: epizootiology in Chinook Salmon of central Oregon. Transactions of the American Fisheries Society 110:507–513.

Ray, R. A., and J. L. Bartholomew. 2013. Estimation of transmission dynamics of the *Ceratomyxa shasta* actinospore to the salmonid host. Parasitology 140:907–916.

Ray, R. A., R. A. Holt, and J. L. Bartholomew. 2012. Relationship between temperature and *Ceratomyxa shasta*-induced mortality In Klamath River salmonids. Journal of Parasitology 98:520–526.

Ray, R. A., P. A. Rossignol, and J. L. Bartholomew. 2010. Mortality threshold for juvenile Chinook Salmon (*Oncorhynchus tshawytscha*) in an epidemiological model of *Ceratomyxa shasta*. Diseases of Aquatic Organisms 93:63–70.

Richter, A., and S. A. Kolmes. 2005. Maximum temperature limits for Chinook, Coho, and Chum salmon, and steelhead trout in the Pacific Northwest. Reviews in Fisheries Science 13:23–49.

Steyerberg, E. W., F. E. Harrell Jr, G. J. J. M. Borsboom, M. J. C. Eijkemans, Y. Vergouwe, and J. D. F. Habbema. 2001. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. Journal of Clinical Epidemiology 54:774–781.

Thomas-Jinu, S., and A. E. Goodwin. 2004. Acute columnaris infection in Channel Catfish, *Ictalurus punctatus* (Rafinesque): efficacy of practical treatments for warmwater aquaculture ponds. Journal of Fish Diseases 27:23–28.

Thompson, K. G., R. B. Nehring, D. C. Bowden, and T. Wygant. 1999. Field exposure of seven species or subspecies of salmonids to *Myxobolus cerebralis* in the Colorado River, Middle Park, Colorado. Journal of Aquatic Animal Health 11:312–329.

Udey, L. R., J. L. Fryer, and K. S. Pilcher. 1975. Relation of water temperature to ceratomyxosis in Rainbow Trout (*Salmon gairdneri*) and Coho Salmon (*Oncorhynchus kisutch*). Journal of the Fisheries Board of Canada 32:1545–1551.

Zhang, J., and Y. Peng. 2007. A new estimation method for the semiparametric accelerated failure time mixture cure model. Statistics in Medicine 26:3157–3171.