Received: 24 January 2020 | Accepted: 27 July 2020

DOI: 10.1111/1365-2664.13744

Journal of Applied Ecology BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

A multi-state occupancy modelling framework for robust estimation of disease prevalence in multi-tissue disease systems

Vratika Chaudhary[1] | Samantha M. Wisely[1,2] | Felipe A. Hernández[3] | James E. Hines[4] | James D. Nichols[5] | Madan K. Oli[1]

[1]Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL, USA

[2]School of Natural Resources and Environment, University of Florida, Gainesville, FL, USA

[3]Instituto de Medicina Preventiva Veterinaria, Facultad de Ciencias Veterinarias, Edificio Federico Saelzer, Valdivia, Chile

[4]U.S. Geological Survey, Patuxent Wildlife Research Center, Beltsville, MD, USA

[5]U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, MD, USA

Correspondence
Vratika Chaudhary
Email: chaudhary.vratika@gmail.com

Handling Editor: Silke Bauer

Abstract

1. Given the public health, economic and conservation implications of zoonotic diseases, their effective surveillance is of paramount importance. The traditional approach to estimating pathogen prevalence as the proportion of infected individuals in the population is biased because it fails to account for imperfect detection. A statistically robust way to reduce bias in prevalence estimates is to obtain repeated samples (or sample many tissues in multi-tissue disease systems) and to apply statistical methods that account for imperfect detection and permit the interdependence of the infection process across multiple tissues.
2. We developed a multi-state occupancy modelling framework which considers two scenarios about the infection process, one where no assumptions about the dependencies among the tissues are made (general), and another where dependence among tissues is not permitted (constrained).
3. We applied this framework to pseudorabies virus (PrV) DNA detection data obtained from whole blood; and oral, nasal and genital mucosa of 510 feral swine *Sus scrofa* during the years 2014–2016 in Florida, USA.
4. The constrained model was better supported by data. PrV prevalence estimates varied among tissues and were higher than the naïve estimates, ranging from to 0.06 (CI: 0.02–0.14) in genital to 0.54 (CI: 0.14–0.82) in nasal tissue. Probability of PrV detection ranged from 0.11 (CI: 0.06–0.18) in nasal to 0.51 (CI: 0.21–0.81) in genital tissue.
5. PrV prevalence was not affected by the age or sex of the animal or the year of sampling, but prevalence increased as drought severity increased.
6. The conditional probability of detecting PrV given infection in at least one tissue type within an individual was highest for nasal tissue, suggesting that nasal is the best tissue to sample for PrV surveillance if only one tissue can be sampled, at least for systems with tissue-specific prevalence and detection probabilities similar to ours.
7. *Synthesis and applications*. We focused on inferences about pathogen prevalence in multi-tissue disease systems, dealing with both nondetection and potential dependencies among tissues in infection status. We found strong evidence of

variation in both prevalence and detection probabilities among tissues. Our results emphasize the importance of sampling multiple tissues and of applying inference methods that account for imperfect detection in the surveillance of systemic diseases. The multi-state modelling framework is broadly applicable to the surveillance of pathogens that infect multiple tissues and can be used even when the infection status of the pathogen in one tissue may depend on the infection status of the pathogen in other tissue(s).

KEYWORDS

disease prevalence estimates, herpes virus, imperfect detection, multi-scale occupancy models, multi-state occupancy models, multi-tissue disease systems, pseudorabies virus, systemic diseases

1 | INTRODUCTION

Zoonotic diseases have received attention in recent decades due to the emergence of pathogens resulting in epidemics and pandemics with substantial implications for public health, global economy and agricultural industry (Daszak, Cunningham, & Hyatt, 2000; Holmes et al., 2019). About 75% of emerging infectious diseases are zoonotic in origin, and wild animals are often their primary reservoirs (Cunningham, 2005). Examples of such pathogens include those causing tuberculosis and brucellosis, with implications for the agricultural industry (James & Rushton, 2002), and Ebola virus and SARS-CoV-2, causing disease outbreaks with grave economic and public health implications (Holmes et al., 2019; Wang, Horby, Hayden, & Gao, 2020). Zoonotic and other wildlife diseases also have substantial conservation consequences because they have been implicated as the cause of population decline of many wildlife species, including the African lion *Panthera leo* (attributed to canine distemper virus; Roelke-Parker et al., 1996) and the African wild dog *Lycaon pictus* (caused by rabies; Alexander & Appel, 1994). Thus, quantitatively rigorous surveillance of wildlife diseases and knowledge of factors affecting their emergence/reemergence are important and sometimes mandated by public health, agricultural or conservation authorities (Alexander, Lewis, Marathe, Eubank, & Blackburn, 2012; Kruse, Kirkemo, & Handeland, 2004; Mastin, van den Bosch, van den Berg, & Parnell, 2019).

Disease surveillance programmes frequently involve the systematic collection, analysis, and interpretation of data and dissemination of information for appropriate actions (Hyatt, Aguirre, Jeggo, & Woods, 2015; Robertson, Nelson, MacNab, & Lawson, 2010). Often, the focus of disease surveillance programmes is to estimate disease or pathogen prevalence (hereafter prevalence), defined as either the number of infected animals in a population at a point in time, or as the probability that a randomly selected animal from a population is infected. Difficulties arise in estimating prevalence when there are false negatives due to imperfect pathogen detection, such that the pathogen is not detected in all infected animals in a population or a sample. When imperfect detection occurs, estimates of prevalence based on raw counts of individuals that test positive are biased, because they fail to distinguish true absence of disease from false negatives (Cooch, Conn, Ellner, Dobson, & Pollock, 2012; Gamble et al., 2019; Jennelle, Cooch, Conroy, & Senar, 2007; Oli, Venkataraman, Klein, Wendland, & Brown, 2006; Tabak, Pedersen, & Miller, 2019). Failure to account for imperfect detection can also lead to incorrect inferences for regression coefficients relating pathogen occurrence to animal-level or environmental factors (e.g. sex, habitat type, season; MacKenzie et al., 2017). Ecologists have long recognized the importance of accounting for imperfect detection in ecological studies (e.g. MacKenzie et al., 2017; Royle, Nichols, & Kery, 2005; Seber, 1982; Thompson & Seber, 1994); however, this recognition is relatively new in disease ecology (Bailey, Reid, Forsman, & Nichols, 2009; Cooch et al., 2012; DiRenzo, Che-Castaldo, Saunders, Grant, & Zipkin, 2019; Greenland, 1996; Jennelle et al., 2007; McClintock et al., 2010; Miller, Brehme, Hines, Nichols, & Fisher, 2012; Nichols, Hollmen, & Grand, 2017).

Disease aetiology is often poorly understood for many pathogens, yet it can have consequences for estimating prevalence. For example, many pathogens infect their hosts through multiple routes, and they can cause systemic diseases where pathogens infect, and are shed through, multiple tissues (hereafter, multi-tissue diseases) with little or no understanding of the sequence of infection (Arzt et al., 2011; Hernández et al., 2018). An animal can be in one of several possible states of infection depending on the type or number of tissues affected. Furthermore, infection of one tissue by the pathogen might depend on the infection status of other tissues. There exist no clear guidelines regarding the definition of infection for systemic diseases or how many tissues are to be sampled; the choice of tissue(s) to be sampled can strongly influence pathogen detection and can lead to underestimation of prevalence if the pathogen is present in tissues other than the one being tested. Furthermore, diagnostic methods used may fail to detect the pathogen even when it is present. A rigorous disease surveillance programme must attempt to address these sources of potential bias.

To account for imperfect detection in studies of species distribution and dynamics, ecologists have developed modelling approaches

collectively known as occupancy models (MacKenzie et al., 2002, 2017). Occupancy modelling uses detection/non-detection data obtained from temporally or spatially replicated surveys and provides approximately unbiased estimates of relevant parameters while accounting for false negatives. The detection/non-detection of species at a given site is conceptually similar to the detection/non-detection of a pathogen in an organism or specific tissue from an organism. This recognition has encouraged the use of occupancy modelling approaches for disease surveillance studies (e.g. Lachish, Gopalaswamy, Knowles, & Sheldon, 2012; Mosher et al., 2019). For systemic diseases, tissues of multiple types from the same animal can be collected and tested multiple times. Such data can be analysed within the multi-scale or multi-state occupancy modelling frameworks to estimate prevalence (Colvin, Peterson, Kent, & Schreck, 2015; MacKenzie et al., 2017; Nichols, Hines, MacKenzie, Seamans, & Gutierrez, 2007); we developed two such models for this purpose.

We demonstrate the application of these models using the Pseudorabies virus (PrV) data from a feral swine *Sus scrofa* population in Florida, USA. PrV is a pathogen of agricultural and conservation concerns. Swine are the only natural reservoir hosts where the infection is usually subclinical, but the virus can spill over from wild to domestic swine and can also infect carnivores via consumption of infected carcasses (see Appendix S1). Intraspecific transmission typically occurs via oro-nasal or sexual contact. In primary infections or during recrudescence, the virus spreads within an animal infecting multiple tissues throughout the body; however, the mechanism and sequence of viral spread within hosts are poorly understood. To address how these uncertainties can influence the PrV prevalence, we collected oral, blood, nasal and genital samples from individual swine and tested for the presence of PrV DNA using quantitative Polymerase Chain Reaction (qPCR). We analysed resulting data using the occupancy models developed here to estimate tissue-specific PrV prevalence and the probability of virus detection. Since qPCR detects virus particles, our prevalence estimates include only the animals that are actively shedding the virus and are likely infectious. We also tested for the potential role of environmental factors and host attributes as modulators of the viral infection in different tissues. Finally, we develop an objective method for selecting tissue type(s) to sample for systemic disease surveillance based on the conditional probability of detecting the pathogen given infection in at least one tissue type within hosts.

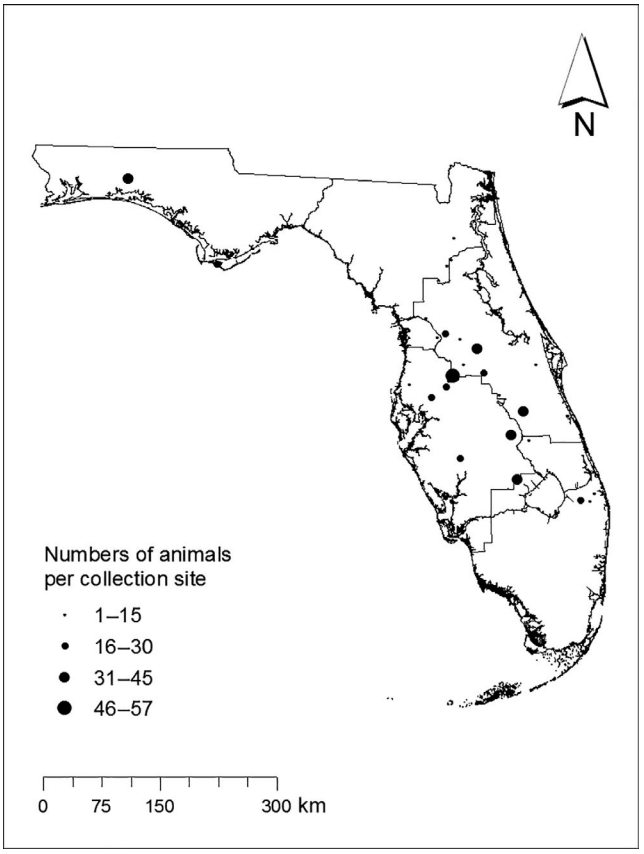

FIGURE 1 Location of feral swine sampled across Florida, USA, from 2014 to 2016. From each sampled animal, blood, oral, nasal and genital tissue samples were collected for molecular testing for the presence of pseudorabies virus

2 | MATERIALS AND METHODS

2.1 | Tissue sample collection and molecular detection of PrV

We opportunistically sampled 549 feral swine at 39 sites across Florida during 2014–2016 (Figure 1) as part of national feral swine disease monitoring effort led by the United States Department of Agriculture. Swine were either euthanized and sampled immediately during animal-control efforts or were hunted by hunters on federal and state wildlife management areas, military bases and private properties and sampled at hunting check-stations. We collected a single sample of whole blood, nasal, oral and genital swabs from the sampled swine; not all tissues were collected from all 549 swine (Table S1). Each sample was tested for PrV DNA up to three times using qPCR. We define 'survey' as at least one qPCR test conducted for the viral presence in at least one of the distinct tissues; thus, our study consisted of as many as three surveys because each tissue was tested for viral presence up to three times. Details regarding feral swine biology, the collection of tissue samples and molecular methods are provided in Appendix S1. Of the 549 swine sampled, age and sex could be determined for 510 swine; only these individuals were used for analyses. We note that false negatives can occur due to sample location (swab taken from a specific location where the virus is absent) or due to diagnostic error; however, we did not decompose the two.

2.2 | Occupancy models for multi-tissue disease systems

Standard occupancy models describe processes occurring at two distinct scales, the distribution of the focal entity (species in original

applications) across sample units (spatial areas in original applications), and the detections or not of the entity across replicate temporal surveys of these sample units (MacKenzie et al., 2002, 2017). In some sampling situations, an intermediate process is inserted, in the form of the distribution of the focal entity across multiple sampling locations within the larger sample unit, with replicate surveys conducted at these locations (MacKenzie et al., 2017; Mordecai, Mattsson, Tzilkowski, & Cooper, 2011; Nichols et al., 2008). This kind of design permits estimation of occupancy at the usual scale of the sample unit, as well as at the local scale of the locations within units. We observe that such multi-scale models are special cases of multi-state models, in which the different patterns of local occupancy (e.g. the number of locations occupied within a sample unit) are viewed as different states of the larger sample unit.

Our use of occupancy modelling to deal with multi-tissue systems appears to be a clear example of multi-scale modelling in which the individual organism is the sample unit and the different tissue types are the locations within sites. Indeed, our simplest model assuming independence of pathogen presence among the different tissue types within an individual is very similar to the multi-scale occupancy models of Nichols et al. (2008) and Mordecai et al. (2011).

In our more general model, we sought to relax this assumption of independence of pathogen presence among tissue types. This assumption may be violated in systemic diseases where infection can occur via multiple routes, and the pathogen can infect various tissues in a particular sequence that is frequently unknown (Arzt et al., 2011; Hernández et al., 2018). Depending on the transmission route, infection of neighbouring tissues may be more likely than those that are spatially separated. Therefore, multi-tissue disease systems present the potential problem of spatially or temporally correlated occupancy of the spatial locations (i.e. tissues or organs) within sample units (animals). To address the complexity of modelling multi-tissue disease systems, we developed a multi-state occupancy modelling framework in which each combination of infected tissue types within an organism is viewed as a different state. For this reason, we refer to our models as multi-state, recognizing that they can be viewed as multi-scale as well. Our approach permits estimation of host- and tissue-specific pathogen prevalence while accounting for imperfect detection that may vary among tissue types. We then use this multi-state occupancy modelling framework to estimate the prevalence of PrV, in feral swine samples collected from Florida, USA. Here we specifically deal with the issue of imperfect detection of the pathogen within a host; however, this approach can be extended to address the issue of imperfect detection of the host as well (see Section 4).

We consider a situation where $s = 1, 2, \ldots, S$ tissues within a host may be infected by a pathogen. At any given time, the pathogen may be present in a tissue (true disease state = 1) or not (true disease state = 0). In our most general model (in the sense of most parameters and fewest assumptions), each potential set of infected tissues in an organism is defined as that organism's disease state. Thus, for S tissues, 2^S states are possible. For example, an animal can be infected in each of the four tissue types sampled (oral, blood, nasal and genital; $S = 4$; true disease state = {1111}), or none is infected (true disease state = {0000}) or a combination of ≥2 tissues are infected; this scenario leads to $2^4 = 16$ possible true disease states (Table 1). Although the presence of viral DNA is indicative of current or recent past infection, we assume that a sample that tested positive for viral DNA contains the virus.

A negative test result can arise either from the true absence of the pathogen in the tissue or because of failure to detect the pathogen that was present. Failure to detect the pathogen can occur due to sampling of the tissue or due to a diagnostic error. Thus, for each replicated survey j (a qPCR assay in our case), a sampled animal i ($i = 1, 2, \ldots, N$) has an associated observation state, $x_{i,j}$. For every survey j, a tissue sampled from animal i can take the value of 1 if the pathogen is detected, 0 otherwise. For example, observation state $x_{i,j}$ = {1111} indicates that the pathogen is detected in all four

TABLE 1 Definition of each of the 16 possible infection states based on the four sampled tissues [oral (O), blood (B), nasal (N) and genital (G)]. Estimates of the proportion of animals in each possible state ($\widehat{\varphi}$) and *SE* obtained from the general multi-state model also are reported. The states are mutually exclusive, and these probabilities sum to 1; thus, for 16 possible disease states, only 15 parameters need to be estimated

States OBNG	Definition	$\widehat{\varphi}$	*SE*
0000	PrV was present in none of the tissues	0.333	0.193
0001	PrV was present only in genital tissue	0.078	0.040
0010	PrV was present only in nasal tissue	0.240	0.206
0011	PrV was present in nasal and genital tissue	0.000	0.000
0100	PrV was present only in blood	0.015	0.176
0101	PrV was present in blood and genital tissue	0.000	0.000
0110	PrV was present in blood and nasal tissue	0.211	0.197
0111	PrV was present in blood, nasal and genital tissue	0.000	0.000
1000	PrV was present only in oral tissue	0.043	0.082
1001	PrV was present in oral and genital tissue	0.000	0.000
1010	PrV was present in oral and nasal tissue	0.044	0.089
1011	PrV was present in oral, nasal and genital tissue	0.000	0.000
1100	PrV was present in oral and blood tissue	0.000	0.000
1101	PrV was present in oral, blood and genital tissue	0.000	0.000
1110	PrV was present in oral, blood and nasal tissue	0.036	0.042
1111	PrV was present in oral, blood, nasal and genital tissue	0.000	0.000

sampled tissues, whereas observation state $x_{i,j} = \{0000\}$ indicates that the pathogen is not detected in any of the four sampled tissues and so on. For true state {1111}, all observation states ($x_{i,j}$) are possible. For all other true states, some observation states are not possible, as we assume false positives do not occur (no false positives due to laboratory procedures or field sample collection). For example, if true state is {0011} (i.e. the first two tissues do not contain the virus), then observation states {0111}, {1011} and {1111} are not possible.

The survey-specific observation state depends on the animal's true infection state, the number of tissues sampled and the tissue-specific detection probabilities. Pathogen detection histories can be modelled using the probabilities associated with each true state, φ^m, and the tissue-specific detection probabilities, $p_j^s = \left(p_1^s, p_2^s, \ldots, p_j^s\right)$. Let φ^m be the probability that the animal is in state m. Let p_j^s be the probability of observing the pathogen in a given infected tissue s (e.g. $s = O$, oral; B, blood; N, nasal; G, genital) in survey j. As an example, consider the detection history (0101, 0011), with pathogen detections in blood and genital tissue during the first survey (assay), and in nasal and genital tissues in the second survey. We know from the detection history that blood, nasal and genital tissues are infected, so the only uncertainty concerns the infection status of oral tissue. We can model this detection history as follows:

$$\begin{aligned} &Pr(h_i = (0101, 0011) \,|\, \varphi^m, p^s) \\ &= \left[\varphi^{1111}\left(1 - p_1^O\right) p_1^B \left(1 - p_1^N\right) p_1^G \left(1 - p_2^O\right) \left(1 - p_2^B\right) p_2^N p_2^G\right] \\ &\quad + \left[\varphi^{0111} p_1^B \left(1 - p_1^N\right) p_1^G \left(1 - p_2^B\right) p_2^N p_2^G\right]. \end{aligned}$$

The probability associated with each possible detection history can be written in the above manner, and the likelihood is simply the product of these probabilities for the N animals that are selected for testing:

$$L(\varphi, p^O p^B p^N p^G | h) = \prod_{i=1}^{N} Pr(h_i), \tag{1}$$

where φ is a vector of the probabilities that the animal is in each state m (φ^m; Table 1). The states are mutually exclusive; therefore, these probabilities sum to 1. An animal can only be in one state at a given time, and one of those states is 'no tissues infected'. Thus, the parameterization of φ provides a very general description of state-specific infection ('general model').

A more restrictive hypothesis is that the probability of any tissue, s, being infected is not dependent on the infection status of any of the other tissues in the same animal ('constrained model'). ψ_i^s is defined as the probability that tissue type s of animal i is infected. Under the hypothesis of no dependence, the probabilities for each of the m true infection states can be written as a function of these four tissue-specific infection probabilities. For example, $\varphi^{1011} = \psi^O(1 - \psi^N)\psi^B\psi^G$. This model requires only four infection parameters, as contrasted with the 15 parameters of the general model (Table 1) that requires no knowledge of the dependencies among the different tissue-specific infection probabilities.

Estimates of the φ^m, or the probabilities comprising φ, estimate the proportion of animals in the sample found in each disease state (Table 1). These prevalence estimates may be of intrinsic interest, and they can be combined to obtain the kinds of prevalence estimates that are more commonly reported. For example, if overall prevalence is defined as the proportion of swine that have at least one of the four tissue types infected by the pathogen (denote as ψ^*), then under the most general model this can be estimated as ($\widehat{\psi}^* = 1 - \widehat{\varphi}^{0000}$), the complement of the estimated probability of no tissues infected. Under the constrained model assuming no dependence of infection among tissue types, overall prevalence can be estimated as: $\widehat{\psi}^* = 1 - (1 - \widehat{\psi}^O)(1 - \widehat{\psi}^B)(1 - \widehat{\psi}^N)(1 - \widehat{\psi}^G)$. In some cases, the tissue-specific prevalence may be of interest. Under the most general model, the probability that an animal has infected blood can be obtained by summing the estimates of φ^m that indicate an infection of blood: $\widehat{\varphi}^{0100} + \widehat{\varphi}^{1100} + \widehat{\varphi}^{0110} + \widehat{\varphi}^{1101} + \widehat{\varphi}^{1110} + \widehat{\varphi}^{1101} + \widehat{\varphi}^{0111} + \widehat{\varphi}^{1111}$. Under the simpler model of no dependencies among tissue infections, the estimated prevalence for blood is $\widehat{\psi}^B$. The central point is that there are multiple ways to define the prevalence, with the definition of choice dependent on the biological questions being addressed. The approach that we have described is sufficiently general to be useful regardless of the definition that is used. We note that our terminology of 'general' and 'constrained' models describes the infection process only. We assume no dependencies among tissue-specific detection probabilities in either model.

In addition to the estimation of tissue-specific prevalence, we tested for the influence of the following covariates on PrV prevalence for both general and constrained models: (a) age of the animal (covariate 'age'; adult, subadult or juvenile); (b) sex of the animal (covariate 'sex'; male or female); (c) year of sampling (covariate 'year'; sampling occurred in 3 years, 2014–2016; and (d) Palmer Drought Severity Index (PDSI), a measure of monthly drought severity (covariate 'PDSI'; Table 2). PDSI estimates relative dryness and ranges from −10 (extreme drought) to +10 (wet) and is standardized to 0. We downloaded monthly PDSI values for years 2014–2016 for the state of Florida from the website of the National Weather Service (https://www.cpc.ncep.noaa.gov/products/Drought/). Specifically, we tested the following hypotheses: PrV prevalence: (a) will be higher in nasal and oral tissues than in genital and blood, as feral swine maintain social bonds by touching snouts, and pathogen transmission through the oro-nasal pathway is most likely; (b) will be greater in females as they are more social than males; (c) will be higher in juveniles and subadults, as adults develop life-long immunity to the virus once infected; and (d) will be higher during dry conditions (we used the PDSI for the month and year in which the sample was obtained) as drought can create a physiologically stressful situation for swine which may lead to viral recrudescence (Tanaka, Imamura, Sakaguchi, Mannen, & Matsuo, 1996).

We used a maximum likelihood approach implemented in the R computing environment (R Core Team, 2019) for model fitting and parameter estimation. The effects of covariates on the parameters, ψ^s, of the constrained model were modelled using the logit link function. The effects of covariates on the parameters, φ^m, of the more general model were modelled using the multinomial logit-link function (e.g. MacKenzie et al., 2017; Nichols et al., 2007). Furthermore, due to logistical constraints, not all tissues were sampled from all

Model type	Model	AIC	ΔAIC	K^a	negLL[b]
Constrained	**ψ (tissue * PDSI) p (tissue)**	**697.34**	**0.00**	**12**	**336.67**
Constrained	ψ (tissue + PDSI) p (tissue)	701.54	4.20	9	341.77
Constrained	ψ (PDSI) p (tissue)	702.48	5.13	6	345.23
Constrained	ψ (tissue + year) p (tissue)	709.87	12.53	10	344.93
Constrained	ψ (year) p (tissue)	712.57	15.23	7	349.28
General	φ (state * PDSI) p (tissue)	713.15	15.81	34	322.57
General	φ (state + PDSI) p (tissue)	714.35	17.01	20	337.17
Constrained	ψ (tissue * year) p (tissue)	715.35	18.00	16	341.67
Constrained	ψ (tissue) p (tissue)	722.12	24.77	8	353.05
Constrained	ψ (tissue + sex) p (tissue)	723.50	26.15	9	352.74
Constrained	ψ (.) p (tissue)	724.97	27.63	5	357.48
Constrained	ψ (tissue + age) p (tissue)	725.96	28.61	10	352.97
Constrained	ψ (sex) p (tissue)	726.45	29.11	6	357.22
General	φ (state + year) p (tissue)	727.89	30.54	21	342.94
Constrained	ψ (age) p (tissue)	728.21	30.87	7	357.10
Constrained	ψ (tissue * sex) p (tissue)	728.42	31.07	12	352.20
Constrained	ψ (tissue * age) p (tissue)	730.12	32.78	16	349.06
General	φ (state) p (tissue)	739.02	41.67	19	350.50
General	φ (state + sex) p (tissue)	739.73	42.38	20	349.86
General	φ (state + age) p (tissue)	741.44	44.10	21	349.72
General	φ (state * sex) p (tissue)	764.79	67.45	34	348.39
General	φ (state * year) p (tissue)	776.61	79.27	49	345.41

TABLE 2 Comparison of all the tested models incorporating additive and interactive effects of covariates (age and sex, year of survey and Palmer Drought Severity Index, PDSI) on pseudorabies virus (PrV) prevalence. Constrained models are based on tissue type, and the probability of any tissue being infected is not dependent on the infection status of any other tissue. General models are based on states defined as combinations of infected tissue types, and no assumptions about the dependencies among the tissues are required. The probability that a particular tissue is infected is denoted ψ (constrained model); the probability that the animal is in a particular state of infection is denoted φ (general model); and p is the probability of PrV detection. A '+' sign indicates an additive effect, and a '*' indicates both additive and interactive effects, 'K' indicates the number of parameters and 'negLL' indicates negative log-likelihood. The model with the lowest AIC value was considered to be the best-supported model (bold font); ΔAIC refers to the difference in AIC value of the top model and the model with which comparison is being made

surveyed animals, resulting in missing values; our model accommodates missing values in such cases. We used an information-theoretic approach based on the Akaike Information Criterion (AIC; Burnham & Anderson, 2002) for model comparison and statistical inference. We used a parametric bootstrapping approach to estimate 95% confidence intervals (CI) for all parameters using 10,000 bootstrap samples (Efron & Tibshirani, 1993).

2.3 | Sampling for systemic disease surveillance

One challenge inherent in systemic disease surveillance is that there is no guidance regarding which or how many tissues are to be sampled. When resources are limited and/or when only one tissue can be sampled, it would be useful to know which tissue type to sample and how many times. One way to approach this question is to compute the conditional probability of detecting infection in at least 1 of K surveys of a specific tissue type, s, given infection in at least one tissue type within an individual (denoted as P^{sK}). For example, in our case with four tissue types, oral (O), blood (B), nasal (N) and genital (G), we can estimate the conditional probability of detecting infection if we only sample blood:

$$\widehat{P}^{BK} = \frac{\widehat{\psi}^{B}\left(1-\left[1-\widehat{p}^{B}\right]^{K}\right)}{1-\left(1-\widehat{\psi}^{O}\right)\left(1-\widehat{\psi}^{B}\right)\left(1-\widehat{\psi}^{N}\right)\left(1-\widehat{\psi}^{G}\right)}. \quad (2)$$

The numerator of (Equation 2) is the product of the estimated probability that blood is infected and the estimated probability of detection in K surveys, given that blood is infected. The denominator of Equation 2 is simply the estimated probability that the organism is infected (in at least one tissue type). The important point to note about expression 2 is that the probability of detecting infection in a specific tissue type is a function of the likelihood that the tissue is infected, given that the organism is infected, the tissue-specific detection probability and the number of surveys or tissue samples. If enough surveys are conducted to drive the detection probability component of expression 2 to approach 1, then the best tissue to sample will always be the one with the largest conditional probability of infection, given organism-level infection. One can then plot the relationship between $\widehat{P}^{sK}$ and K, and determine the best tissue(s) to sample (Figure 4).

Expression 2 is based on the parameterization of our constrained model, but a similar expression is readily derived for the more general model as well. For example, when blood is given by the second position in the four-entry infection state vector, we can compute $\widehat{P}^{BK}$ as:

$$\widehat{P}^{BK} = \frac{\left(\widehat{\varphi}^{0100}+\widehat{\varphi}^{1100}+\widehat{\varphi}^{0110}+\widehat{\varphi}^{1101}+\widehat{\varphi}^{1110}+\widehat{\varphi}^{1101}+\widehat{\varphi}^{0111}+\widehat{\varphi}^{1111}\right)\left(1-\left[1-\widehat{p}^{B}\right]^{K}\right)}{1-\widehat{\varphi}^{0000}}. \quad (3)$$

3 | RESULTS

Out of 510 swine (237 males and 273 females; 424 adults, 47 subadults and 39 juveniles) sampled, we conducted qPCR on oral, whole blood, nasal and genital (only females) samples collected from 408, 439, 497 and 196 animals respectively. PrV was detected in 14.5% (N = 510) of the animals in at least one of the four tissues that was tested; it was detected in 3.1% of oral (N = 408), 6.6% of blood (N = 439), 6.0% of nasal (N = 497) and 4.0% of genital tissue (N = 196) samples. PrV was detected infrequently in more than one tissue within an individual animal (Table S1).

We analysed detection history data using both the general and constrained multi-state occupancy models. The constrained models were better supported by data, as evidenced by the lower AIC values, with little or no evidence of dependencies among tissues of PrV infection (Table 2). We present the estimates under the general model with no covariates, simply to illustrate the different nature of the resulting estimates (Table 1, Figure 2a). The constrained model ψ (tissue * PDSI) p (tissue) was the best-supported model. Based on this model, PrV prevalence was 0.12 (CI: 0.04–0.29) in oral tissue, 0.24 (CI: 0.10–0.39) in blood, 0.54 (CI: 0.14–0.82) in nasal and 0.06 (CI, 0.02–0.14) in genital tissue (Figure 2a). The probability of PrV detection was lowest in nasal tissue (0.11, CI: 0.06–0.18) and highest in genital tissue (0.51, CI: 0.21–0.81; Figure 2b). Naïve estimates of PrV prevalence were substantially lower than those produced by the occupancy models (Figure 2a). The estimated overall PrV prevalence in our sample was 0.71 (CI: 0.35, 1), which is about five times higher than the naïve occupancy probability (0.14).

PDSI was negatively related to PrV prevalence in all except oral tissue (Figure 3). We note that the higher prevalence in blood and nasal tissue produced larger sample sizes of infected animals and a greater ability to detect covariate relationships. There was little evidence for the effects of sex or age on prevalence (Table 2), although point estimates for effect parameters in models that included these covariates showed the predicted signs, with prevalence slightly greater for females than males, and for juveniles and subadults than adults (Appendix S2; Figure S2).

The conditional probability of detecting PrV given infection in at least one tissue type within an individual ($\widehat{P}^{sK}$) was highest for nasal tissue and lowest for genital tissue (Figure 4). Also, $\widehat{P}^{sK}$ for nasal tissue increased more rapidly as the number of surveys increased, suggesting that nasal is the best tissue to sample for PrV surveillance if only one tissue can be sampled. We emphasize that the plots of Figure 4 and resulting inferences are conditional on the estimates of tissue-specific prevalence and detection probabilities estimated from our specific sample.

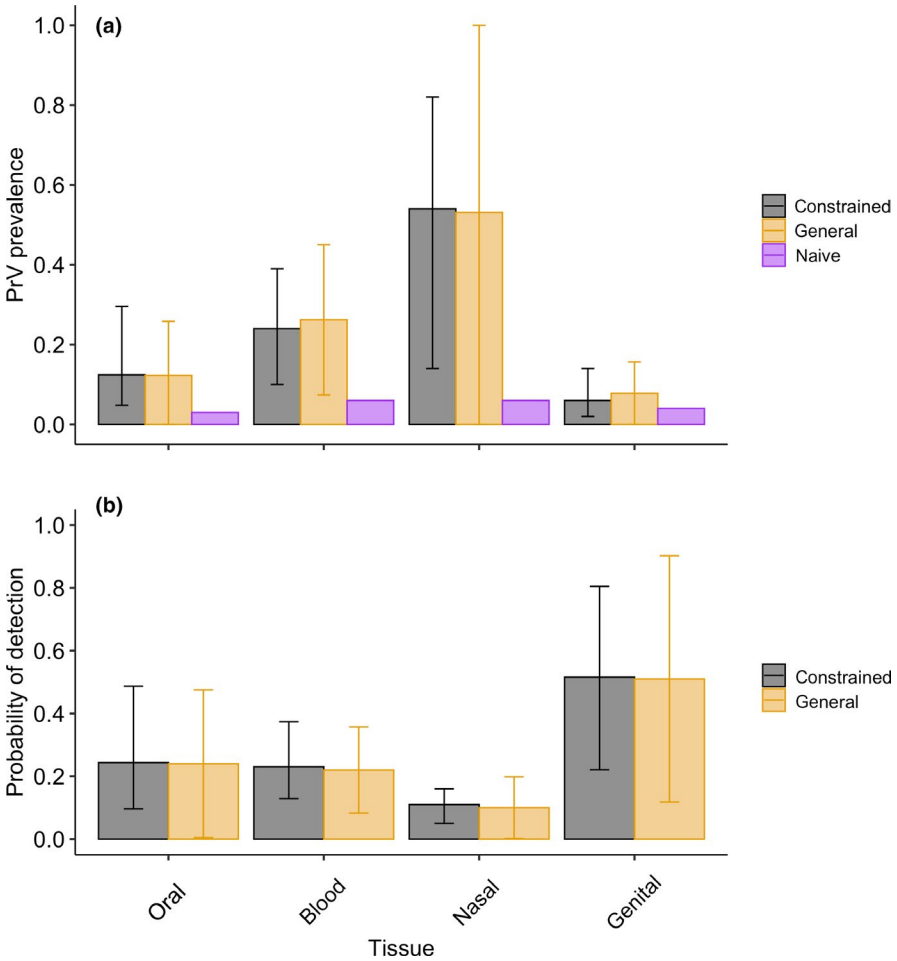

FIGURE 2 (a) Pseudorabies virus (PrV) prevalence estimates ($\widehat{\psi}$) and associated 95% confidence intervals (CI) for each tissue based on the best-supported model (Table 2; constrained model ψ (tissue * PDSI) p (tissue)), with the value of Palmer Drought Severity Index (PDSI) fixed to zero. Tissue-specific PrV prevalence estimates based on general φ (states) p (tissue) model are also reported to illustrate the different nature of the resulting estimates. Naïve estimates of occupancy probability (proportion of positive samples out of total samples tested) also are presented. (b) Probability of detection of PrV (along with CI) for each tissue type based on the constrained model ψ (tissue * PDSI) p (tissue) and general φ (states) p (tissue) model

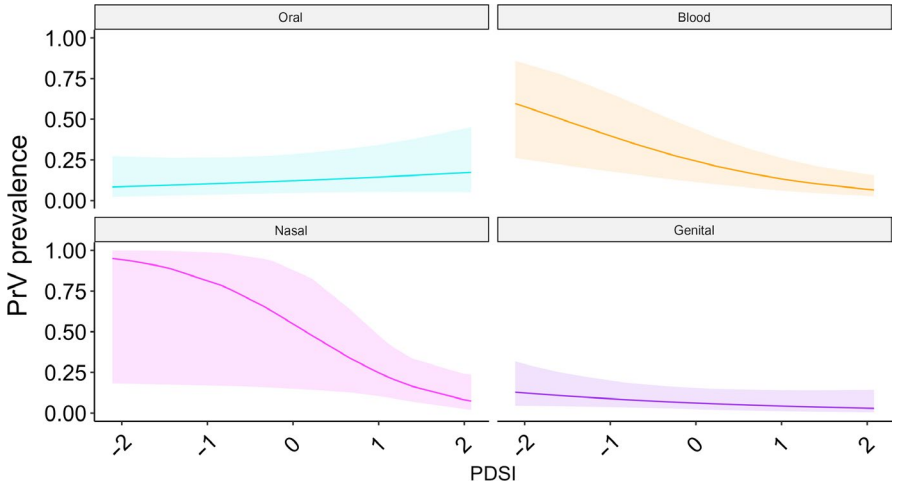

FIGURE 3 Pseudorabies virus (PrV) prevalence estimates ($\widehat{\psi}$) and associated 95% confidence intervals (CI) showing the effect of Palmer Drought Severity Index (PDSI) on ψ, based on the ψ (tissue * PDSI) p (tissue) model

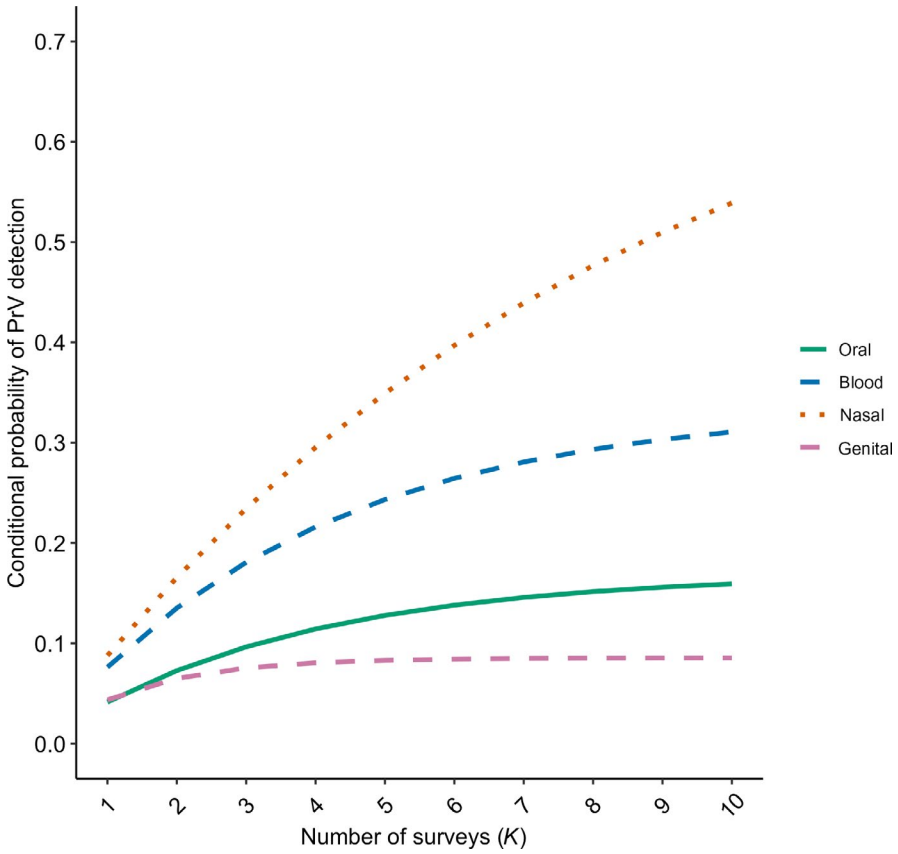

FIGURE 4 Conditional probability of detecting pseudorabies virus (PrV) in a tissue (oral, blood, nasal or genital) given the individual animal is infected (P^{sK}), as a function of the number of surveys (K; replicate qPCR assays in our case), calculated using expression (2)

4 | DISCUSSION

In this study, we developed a multi-state occupancy modelling framework to address some of the challenges inherent in estimating prevalence in multi-tissue disease systems. Our model permits estimation of tissue-specific infection, as well as overall pathogen prevalence, under two different hypotheses about tissue-infection dependencies. The simpler hypothesis is that the infection status of any single tissue type does not depend on the infection status of another tissue type (constrained model). The more general hypothesis permits the infection status of one tissue type to be dependent on the status of another tissue type (general model). Models based on either of these hypotheses simultaneously account for imperfect detection to provide approximately unbiased estimates of pathogen prevalence and allow for modelling the effects of relevant covariates to understand their roles as predictors of infection status. We used this modelling approach to estimate PrV prevalence in feral swine populations of Florida and to test for factors affecting the disease prevalence.

Even though the sequence of PrV dissemination within the host body is not completely understood, the oro-nasal route is believed to be an important route of direct transmission (Pirtle, Sacks, Nettles, & Rollor, 1989; Verin, Varuzza, Mazzei, & Poli, 2014).

Therefore, we expected that PrV prevalence would be highest in nasal and oral tissue and lowest in genital tissue, assuming little or no sexual transmission occurs. This is because by the time virus is transported to genital tissue, the immune system of the animal should be activated and viral load should decrease with time as is the case with other herpes viruses (von Hofsten, Bergstrom, & Zetterberg, 2019). Consistent with these expectations, we found that PrV prevalence in genital tissue was lowest (0.06, CI = 0.02–0.14). Although the virus can be sexually transmitted (Romero et al., 2001), the low prevalence of PrV in genital tissue (Figure 2a) suggests that sexual transmission is less common than oro-nasal transmission in our study population. We also found that PrV prevalence was highest in nasal tissue (0.51, CI = 0.21–0.81). A consequence of low detection probability is that very few nasal samples tested positive for the PrV in more than one survey. Although the mean quantity of extracted DNA from nasal samples, as well as the mean quantification cycle values, were similar to those of other tissues, template DNA may have contained more PCR inhibitors in samples taken from nasal passages which can interfere with the PCR process. Naïve estimates of PrV infection in feral swine were 1.5–9 times lower than prevalence estimates that accounted for imperfect detection, suggesting that failure to account for imperfect detection would have led to substantial underestimation of PrV prevalence (Figure 2a).

PrV prevalence varied over time due to variation in drought conditions as indicated by the influence of the PDSI. Limitation of water during dry periods can create stressful physiological conditions making swine vulnerable to infection or increasing the chances of viral recrudescence (Tanaka et al., 1996). Our results show a strong negative effect of drought conditions on PrV prevalence, specifically for nasal and blood tissue, suggesting that the feral swine may pose a greater risk of infecting other feral swine and domestic swine during dry periods.

Estimates of systemic disease prevalence depend both on the number of tissues or organs to be sampled, and the number of replicate samples per tissue type. Because resources are rarely unlimited, researchers would have to find a balance between how many tissues to sample and how many times. We have proposed an objective approach to address this question based on the conditional probability of detecting infection in at least 1 of K surveys of a specific tissue type, s, given infection in at least one tissue type within an individual (P^{sK}; expressions 2–3). In our study system, $\widehat{P}^{sK}$ for nasal tissue was generally higher than those for other tissue types, suggesting that, for our sampled population, nasal is the best tissue to sample for PrV surveillance if only one tissue can be sampled (Figure 4). However, note that this result is entirely a function of the high prevalence of the virus in this tissue type. If tissue-specific prevalence is different in a different system, this result may not hold. The cost can also be taken into account if there is variation among tissues in the cost per sample. For example, if the analysis of one tissue type (A) is twice as expensive as that of another tissue type (B), then the cost of tissue A for $K = 1$ sample is equivalent to that of tissue B for $K = 2$ samples. The probability of detection per unit cost can be computed for each tissue type in this manner.

The modelling framework presented here was developed to deal with the issue of imperfect detection of the pathogen, but infected individuals are often detected (and thus appear in samples for prevalence calculations) with different probabilities than uninfected individuals (e.g. Conn & Cooch, 2009; Cooch et al., 2012; Jennelle et al., 2007). Our samples came entirely from animals killed during sport hunting operations or those that were euthanized for animal control. If infected animals are less wary or perhaps slower than uninfected animals, then we would expect them to appear in our sample with a higher probability than uninfected animals. The individual prevalence and tissue-specific prevalences that we report above strictly correspond to our opportunistic sample of 510 individuals. However, we can use multistate capture–recapture or removal modelling (Arnason, 1973; Brownie, Hines, Nichols, Pollock, & Hestbeck, 1993; Lebreton, Nichols, Barker, Pradel, & Spendelow, 2009) to model variation in animal-level detection probability when estimating prevalence at the population level. We can then directly estimate the population-level prevalence by developing a joint likelihood that incorporates both animal-level and pathogen-level sampling processes.

For our data, the constrained model was better supported than the general model, likely because there were very few cases of concurrent PrV shedding in more than one tissue (Table 2). This might not be the case for other pathogens, and it is advisable to fit general and constrained models and use information-theoretic (Burnham & Anderson, 2002) or other approaches to select the model most appropriate for the data. This approach would be particularly valuable for the surveillance of other little-known pathogens that exhibit pantropism. For example, the Zika virus can infect multiple organs (Miner & Diamond, 2017), but little is known about the dependence among tissues during the infection process. In such cases, our general model permits state-specific estimates of prevalence without having to make assumptions regarding the interdependence of infection among tissues (or lack thereof). When an understanding of disease progression is of interest, specific states of infection in the form of various tissue combinations would be expected to occur more or less frequently than expected under a null hypothesis of no-dependence. Such hypotheses can be tested, and a progression path can be proposed and then evaluated using model selection methods within our modelling framework.

Our results suggest that measures of prevalence should be carefully defined, and tissue(s) to be sampled and tested carefully chosen based on any knowledge of tissue-specific infection. This is particularly important for pathogens that infect multiple tissues and/or when infections are subclinical (e.g. *Mycobacterium tuberculosis* causing pulmonary and extrapulmonary TB). We would have grossly underestimated PrV prevalence in our sample of feral swine if these estimates were based on genital tissues only. Depending on the definition of prevalence (e.g. infection of one tissue, or combination of ≥2 tissues), estimates of PrV prevalence in our sample of feral swine would range from <0.001 to 0.54

(Table 1; Figure 2a). As qPCR only detects pathogens in animals that are actively shedding, many more swine might be carrying the latent infection in our population, and those animals were not accounted for in our study.

It is conceivable that the detection of a pathogen in a sample is dependent on the viral load of the sampled tissue (local pathogen abundance). Typically, local pathogen abundance and the probability of detection are positively related such that when viral load is low, the probability of detection is also low (Royle & Nichols, 2003). PrV detection by qPCR is dependent on a threshold viral load (Ren et al., 2018), and it is possible that viral presence in samples collected from tissues with low viral load might have been missed by qPCR assays. Unfortunately, we could not examine the relationship between viral load and the probability of viral detection due to data limitations.

The occupancy modelling framework presented here represents an important advance in disease surveillance of multi-tissue disease systems because it: (a) permits estimation of, and inference about, multi-tissue disease prevalence parameters without assumptions about dependence (or lack thereof) of infection among tissues; (b) offers a flexible framework regarding the definition of what constitutes an infection, depending on the disease biology and focal question(s); (c) appropriately accounts for imperfect detection of the pathogen; and (d) allows modelling of state-specific parameters as functions of covariates.

This modelling approach is conceptually analogous to that used in the modelling of interspecific interactions (MacKenzie, Bailey, & Nichols, 2004; MacKenzie et al., 2017; Miller et al., 2012; Yackulic et al., 2014). The different tissue types in our application are similar to the different species, and the constrained model is analogous to a model of no interspecific interactions. The species interaction factor used in two-species studies estimates the degree to which the probability of finding both species together is more or less than expected under a hypothesis of complete independence (MacKenzie et al., 2004, 2017). Using our general model, we could similarly ask about the degree of departure from independent prevalence for various combinations of tissue types. Extensions to dynamic multi-state models allow investigators to estimate parameters associated with transitioning from one state to another (e.g. Miller et al., 2012; Yackulic et al., 2014). A repeated non-lethal sampling of individual swine over time could be used to investigate whether the probability of one tissue type becoming infected at time $t + 1$ depends on the types of tissues infected at time t, for example. We believe that both single-season and dynamic multi-state occupancy modelling hold great promise for future disease modelling.

ACKNOWLEDGEMENTS

We thank the Department of Wildlife Ecology and Conservation, the Center for Tropical Conservation and Development, (University of Florida), University of Florida Biodiversity Institute and, WildLandscapes International, for supporting VC's graduate studies. F.A.H. contributed to the manuscript while being funded by ANID/CONICYT Fondecyt Postdoctorado number 3180111. We are grateful to E. Cooch, M. Nunez, S. Bauer and an anonymous reviewer for many helpful comments and suggestions, which improved the quality of this paper. We thank the United States Department of Agriculture and Florida Wildlife Services field personnel and hunting check-station operators for help in collecting feral swine samples. Any use of trade, product or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

AUTHORS' CONTRIBUTIONS

V.C., M.K.O., S.M.W. and J.D.N. conceived the ideas and designed methodology; F.A.H. and S.M.W. collected the data; V.C. and J.E.H. analysed the data; V.C., J.D.N. and M.K.O. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT

Data are available via the Dryad Digital Repository https://doi.org/10.5061/dryad.z8w9ghx94 (Chaudhary et al., 2020).

ORCID

Vratika Chaudhary https://orcid.org/0000-0001-7155-122X
Madan K. Oli https://orcid.org/0000-0001-6944-0061

REFERENCES

Alexander, K. A., & Appel, M. J. G. (1994). African wild dogs (*Lycaon pictus*) endangered by a canine-distemper epizootic among domestic dogs near the Masai-Mara National Reserve, Kenya. *Journal of Wildlife Diseases*, *30*, 481–485. https://doi.org/10.7589/0090-3558-30.4.481

Alexander, K. A., Lewis, B. L., Marathe, M., Eubank, S., & Blackburn, J. K. (2012). Modelling of wildlife-associated zoonoses: Applications and caveats. *Vector-Borne and Zoonotic Diseases*, *12*, 1005–1018.

Arnason, A. N. (1973). The estimation of population size, migration rates, and survival in stratified population. *Researches in Population Ecology*, *15*, 1–8.

Arzt, J., Baxt, B., Grubman, M. J., Jackson, T., Juleff, N., Rhyan, J., … Rodriguez, L. L. (2011). The pathogenesis of foot-and-mouth disease II: Viral pathways in swine, small ruminants, and wildlife; myotropism, chronic syndromes, and molecular virus-host interactions. *Transboundary and Emerging Diseases*, *58*, 305–326. https://doi.org/10.1111/j.1865-1682.2011.01236.x

Bailey, L. L., Reid, J. A., Forsman, E. D., & Nichols, J. D. (2009). Modelling co-occurrence of northern spotted and barred owls: Accounting for detection probability differences. *Biological Conservation*, *142*, 2983–2989.

Brownie, C., Hines, J. E., Nichols, J. D., Pollock, K. H., & Hestbeck, J. B. (1993). Capture-recapture studies for multiple strata including non-Markovian transitions. *Biometrics*, *49*, 1173–1187. https://doi.org/10.2307/2532259

Burnham, K., & Anderson, D. (2002). *Model selection and multi-model inference* (2nd ed.). New York, NY: Springer.

Chaudhary, V., Wisely, S. M., Hernandez, F. A., Hines, J. E., Nichols, J. D., & Oli, M. K. (2020). Data from: A multi-state occupancy modeling framework for robust estimation of disease prevalence in multi-tissue disease systems. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.z8w9ghx94

Colvin, M. E., Peterson, J. T., Kent, M. L., & Schreck, C. B. (2015). Occupancy modelling for improved accuracy and understanding

of pathogen prevalence and dynamics. *PLoS ONE*, *10*, e0116605. https://doi.org/10.1371/journal.pone.0116605

Conn, P. B., & Cooch, E. G. (2009). Multistate capture-recapture analysis under imperfect state observation: An application to disease models. *Journal of Applied Ecology*, *46*, 486–492. https://doi.org/10.1111/j.1365-2664.2008.01597.x

Cooch, E. G., Conn, P. B., Ellner, S. P., Dobson, A. P., & Pollock, K. H. (2012). Disease dynamics in wild populations: Modelling and estimation: A review. *Journal of Ornithology*, *152*, S485–S509.

Cunningham, A. A. (2005). A walk on the wild side-emerging wildlife diseases – They increasingly threaten human and animal health. *British Medical Journal*, *331*, 1214–1215. https://doi.org/10.1136/bmj.331.7527.1214

Daszak, P., Cunningham, A. A., & Hyatt, A. D. (2000). Emerging infectious diseases of wildlife – Threats to biodiversity and human health. *Science*, *287*, 443–449. https://doi.org/10.1126/science.287.5452.443

DiRenzo, G. V., Che-Castaldo, C., Saunders, S. P., Grant, E. H. C., & Zipkin, E. F. (2019). Disease-structured N-mixture models: A practical guide to model disease dynamics using count data. *Ecology and Evolution*, *9*, 899–909.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman & Hall.

Gamble, A., Ramos, R., Parra-Torres, Y., Mercier, A., Galal, L., Pearce-Duvet, J., … Boulinier, T. (2019). Exposure of yellow-legged gulls to *Toxoplasma gondii* along the Western Mediterranean coasts: Tales from a sentinel. *International Journal for Parasitology: Parasites and Wildlife*, *8*, 221–228. https://doi.org/10.1016/j.ijppaw.2019.01.002

Greenland, S. (1996). Basic methods for sensitivity analysis of biases. *International Journal of Epidemiology*, *25*, 1107–1116.

Hernández, F. A., Sayler, K. A., Bounds, C., Milleson, M. P., Carr, A. N., & Wisely, S. M. (2018). Evidence of pseudorabies virus shedding in feral swine (*Sus scrofa*) populations of Florida, USA. *Journal of Wildlife Diseases*, *54*, 45–53. https://doi.org/10.7589/2017-04-071

Holmes, J. P., Duff, J. P., Barlow, A., Everest, D., Man, C., Smith, F., & Twomey, F. (2019). 20 years of national wildlife disease surveillance. *Veterinary Record*, *184*, 520–521. https://doi.org/10.1136/vr.l1903

Hyatt, A., Aguirre, A. A., Jeggo, M., & Woods, R. (2015). Effective coordination and management of emerging infectious diseases in wildlife populations. *EcoHealth*, *12*, 408–411. https://doi.org/10.1007/s10393-015-1045-0

James, A. D., & Rushton, J. (2002). The economics of foot and mouth disease. *Revue Scientifique et Technique: Office International des Epizooties*, *21*, 637–644. https://doi.org/10.20506/rst.21.3.1356

Jennelle, C. S., Cooch, E. G., Conroy, M. J., & Senar, J. C. (2007). State-specific detection probabilities and disease prevalence. *Ecological Applications*, *17*, 154–167. https://doi.org/10.1890/1051-0761(2007)017[0154:SDPADP]2.0.CO;2

Kruse, H., Kirkemo, A. M., & Handeland, K. (2004). Wildlife as source of zoonotic infections. *Emerging Infectious Diseases*, *10*, 2067–2072. https://doi.org/10.3201/eid1012.040707

Lachish, S., Gopalaswamy, A. M., Knowles, S. C. L., & Sheldon, B. C. (2012). Site-occupancy modelling as a novel framework for assessing test sensitivity and estimating wildlife disease prevalence from imperfect diagnostic tests. *Methods in Ecology and Evolution*, *3*, 339–348. https://doi.org/10.1111/j.2041-210X.2011.00156.x

Lebreton, J. D., Nichols, J. D., Barker, R. J., Pradel, R., & Spendelow, J. A. (2009). Modelling individual animal histories with multistate capture-recapture models. *Advances in Ecological Research*, *41*, 87–173.

MacKenzie, D. I., Bailey, L. L., & Nichols, J. D. (2004). Investigating species co-occurrence patterns when species are detected imperfectly. *Journal of Animal Ecology*, *73*, 546–555. https://doi.org/10.1111/j.0021-8790.2004.00828.x

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, *83*, 2248–2255. https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2017). *Occupancy estimation and modelling. Inferring patterns and dynamics of species occurrence* (2nd ed.). London, UK: Academic Press.

Mastin, A. J., van den Bosch, F., van den Berg, F., & Parnell, S. R. (2019). Quantifying the hidden costs of imperfect detection for early detection surveillance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*. https://doi.org/10.1098/rstb.2019.0038

McClintock, B. T., Nichols, J. D., Bailey, L. L., MacKenzie, D. I., Kendall, W. L., & Franklin, A. B. (2010). Seeking a second opinion: Uncertainty in disease ecology. *Ecology Letters*, *13*, 659–674. https://doi.org/10.1111/j.1461-0248.2010.01472.x

Miller, D. A. W., Brehme, C. S., Hines, J. E., Nichols, J. D., & Fisher, R. N. (2012). Joint estimation of habitat dynamics and species interactions: Disturbance reduces co-occurrence of non-native predators with an endangered toad. *Journal of Animal Ecology*, *81*, 1288–1297. https://doi.org/10.1111/j.1365-2656.2012.02001.x

Miner, J. J., & Diamond, M. S. (2017). Zika virus pathogenesis and tissue tropism. *Cell Host & Microbe*, *21*, 134–142. https://doi.org/10.1016/j.chom.2017.01.004

Mordecai, R. S., Mattsson, B. J., Tzilkowski, C. J., & Cooper, R. J. (2011). Addressing challenges when studying mobile or episodic species: Hierarchical Bayes estimation of occupancy and use. *Journal of Applied Ecology*, *48*, 56–66. https://doi.org/10.1111/j.1365-2664.2010.01921.x

Mosher, B. A., Brand, A. B., Wiewel, A. N., Miller, D. A., Gray, M. J., Miller, D. L., & Grant, E. H. C. (2019). Estimating occurrence, prevalence, and detection of amphibian pathogens: Insights from occupancy models. *Journal of Wildlife Diseases*, *55*, 563–575. https://doi.org/10.7589/2018-02-042

Nichols, J. D., Bailey, L. L., O'Connell Jr., A. F., Talancy, N. W., Campbell Grant, E. H., Gilbert, A. T., … Hines, J. E. (2008). Multi-scale occupancy estimation and modelling using multiple detection methods. *Journal of Applied Ecology*, *45*, 1321–1329. https://doi.org/10.1111/j.1365-2664.2008.01509.x

Nichols, J. D., Hines, J. E., MacKenzie, D. I., Seamans, M. E., & Gutierrez, R. J. (2007). Occupancy estimation and modelling with multiple states and state uncertainty. *Ecology*, *88*, 1395–1400.

Nichols, J. D., Hollmen, T. E., & Grand, J. B. (2017). Monitoring for the management of disease risk in animal translocation programmes. *EcoHealth*, *14*, S156–S166. https://doi.org/10.1007/s10393-015-1094-4

Oli, M. K., Venkataraman, M., Klein, P., Wendland, L. D., & Brown, M. B. (2006). Population dynamics of infectious diseases: A discrete time model. *Ecological Modelling*, *198*, 183–194.

Pirtle, E. C., Sacks, J. M., Nettles, V. F., & Rollor, E. A. (1989). Prevalence and transmission of pseudorabies virus in an isolated population of feral swine. *Journal of Wildlife Diseases*, *25*, 605–607. https://doi.org/10.7589/0090-3558-25.4.605

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Ren, M., Lin, H., Chen, S., Yang, M., An, W., Wang, Y., … Hu, J. (2018). Detection of pseudorabies virus by duplex droplet digital PCR assay. *Journal of Veterinary Diagnostic Investigation*, *30*, 105–112. https://doi.org/10.1177/1040638717743281

Robertson, C., Nelson, T. A., MacNab, Y. C., & Lawson, A. B. (2010). Review of methods for space-time disease surveillance. *Spatial and Spatio-temporal Epidemiology*, *1*, 105–116. https://doi.org/10.1016/j.sste.2009.12.001

Roelke-Parker, M. E., Munson, L., Packer, C., Kock, R., Cleaveland, S., Carpenter, M., … Appel, M. J. G. (1996). A canine distemper virus

epidemic in Serengeti lions (*Panthera leo*). *Nature*, *379*, 441–445. https://doi.org/10.1038/379441a0

Romero, C. H., Meade, P. N., Shultz, J. E., Chung, H. Y., Gibbs, E. P., Hahn, E. C., & Lollis, G. (2001). Venereal transmission of pseudorabies viruses indigenous to feral swine. *Journal of Wildlife Diseases*, *37*, 289–296. https://doi.org/10.7589/0090-3558-37.2.289

Royle, J. A., & Nichols, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology*, *84*, 777–790. https://doi.org/10.1890/0012-9658(2003)084[0777:EAFRPA]2.0.CO;2

Royle, J. A., Nichols, J. D., & Kery, M. (2005). Modelling occurrence and abundance of species when detection is imperfect. *Oikos*, *110*, 353–359. https://doi.org/10.1111/j.0030-1299.2005.13534.x

Seber, G. A. F. (1982). *The estimation of animal abundance and related parameters* (2nd ed.). New York, NY: Macmillan.

Tabak, M. A., Pedersen, K., & Miller, R. S. (2019). Detection error influences both temporal seroprevalence predictions and risk factors associations in wildlife disease models. *Ecology and Evolution*, *9*, 10404–10414. https://doi.org/10.1002/ece3.5558

Tanaka, S., Imamura, T., Sakaguchi, M., Mannen, K., & Matsuo, K. (1996). Acetylcholine activates latent pseudorabies virus in pigs. *Archives of Virology*, *141*, 161–166. https://doi.org/10.1007/BF01718597

Thompson, S. K., & Seber, G. A. F. (1994). Detectability in conventional and adaptive sampling. *Biometrics*, *50*, 712–724. https://doi.org/10.2307/2532785

Verin, R., Varuzza, P., Mazzei, M., & Poli, A. (2014). Serologic, molecular, and pathologic survey of pseudorabies virus infection in hunted wild boars (*Sus scrofa*) in Italy. *Journal of Wildlife Diseases*, *50*, 559–565.

von Hofsten, J., Bergstrom, T., & Zetterberg, M. (2019). Alpha herpes virus type and viral load in intraocular fluids in patients with acute retinal necrosis. *BMJ Open Ophthalmology*, *4*, e000247. https://doi.org/10.1136/bmjophth-2018-000247

Wang, C., Horby, P. W., Hayden, F. G., & Gao, G. F. (2020). A novel coronavirus outbreak of global health concern. *The Lancet*, *395*, 470–473. https://doi.org/10.1016/S0140-6736(20)30185-9

Yackulic, C. B., Reid, J., Nichols, J. D., Hines, J. E., Davis, R., & Forsman, E. (2014). The roles of competition and habitat in the dynamics of populations and species distributions. *Ecology*, *95*, 265–279. https://doi.org/10.1890/13-0012.1

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Chaudhary V, Wisely SM, Hernández FA, Hines JE, Nichols JD, Oli MK. A multi-state occupancy modelling framework for robust estimation of disease prevalence in multi-tissue disease systems. *J Appl Ecol*. 2020;00:1–12. https://doi.org/10.1111/1365-2664.13744