Received: 6 April 2017 | Accepted: 30 July 2017

DOI: 10.1111/2041-210X.12868

RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Imperfect pathogen detection from non-invasive skin swabs biases disease inference

Graziella V. DiRenzo[1] | Evan H. Campbell Grant[2] | Ana V. Longo[1] | Christian Che-Castaldo[3] | Kelly R. Zamudio[4] | Karen R. Lips[1]

[1]Department of Biology, University of Maryland, College Park, MD, USA

[2]U.S. Geological Survey, Patuxent Wildlife Research Center, SO Conte Anadromous Fish Research Lab, Turners Falls, MA, USA

[3]Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY, USA

[4]Department of Ecology & Evolutionary Biology, Cornell University, Ithaca, NY, USA

Correspondence
Graziella V. DiRenzo
Email: grace.direnzo@gmail.com

Present address
Graziella V. DiRenzo, University of California, Santa Barbara, CA, USA

Funding information
National Science Foundation Graduate Research Fellowship; University of Maryland Latin-American Studies office; National Science Foundation

Handling Editor: Louise Johnson

Abstract

1. Conservation managers rely on accurate estimates of disease parameters, such as pathogen prevalence and infection intensity, to assess disease status of a host population. However, these disease metrics may be biased if low-level infection intensities are missed by sampling methods or laboratory diagnostic tests. These false negatives underestimate pathogen prevalence and overestimate mean infection intensity of infected individuals.
2. Our objectives were two-fold. First, we quantified false negative error rates of *Batrachochytrium dendrobatidis* (*Bd*) on non-invasive skin swabs collected from an amphibian community in El Copé, Panama. We swabbed amphibians twice in sequence, and we used a recently developed hierarchical Bayesian estimator to assess disease status of the population. Second, we developed a novel hierarchical Bayesian model to simultaneously account for imperfect pathogen detection from field sampling and laboratory diagnostic testing. We evaluated the performance of the model, using simulations and varying sampling design to quantify the magnitude of bias in estimates of pathogen prevalence and infection intensity.
3. We show that *Bd* detection probability from skin swabs was related to host infection intensity, where *Bd* infections <10 zoospores have <95% probability of being detected. If imperfect *Bd* detection was not considered, then *Bd* prevalence was underestimated by as much as 71%. In the *Bd*-amphibian system, this indicates a need to correct for imperfect pathogen detection in enzootic host populations persisting with low-level infections. More generally, our results have implications for study designs in other disease systems, particularly those with similar objectives, biology, and sampling decisions.
4. Uncertainty in pathogen detection is an inherent property of most sampling protocols and diagnostic tests, where the magnitude of bias depends on the study system, type of infection, and false negative error rates. Given that it may be difficult to know this information in advance, we advocate that the most cautious approach is to assume all errors are possible and to accommodate them by adjusting sampling designs. The modelling framework presented here improves the accuracy in estimating pathogen prevalence and infection intensity.

KEYWORDS
amphibians, *Batrachochytrium dendrobatidis*, chytrid, hierarchical Bayesian model, occupancy model, Panama

1 | INTRODUCTION

Epidemiologists and wildlife managers rely on accurate estimates of disease parameters, such as pathogen prevalence and infection intensity, to assess the risk of disease emergence in wild host populations (e.g. Langwig et al., 2015). Traditionally, disease ecologists have recognised that imperfect host detection (i.e., false negatives) affects the inferences made on disease dynamics, leading to the adoption of capture-mark-recapture methods to correct for imperfect host detection (e.g., Cooch, Conn, Ellner, Dobson, & Pollock, 2011). More recently, however, there has been growing awareness that imperfect pathogen detection biases the estimation of pathogen prevalence and infection intensity (e.g. Lachish, Gopalaswamy, Knowles, & Sheldon, 2012; Miller, Talley, Lips, & Campbell Grant, 2012). Pathogen prevalence tends to be underestimated, whereas mean infection intensity is overestimated when sampling methods or diagnostic tests miss low-level pathogen infections, causing the misclassification of infected hosts as uninfected (e.g. Lachish et al., 2012; Miller et al., 2012). This growing awareness has led to new quantitative methods that provide a platform to correct for disease state misclassification, improving the quality of inference by reducing bias (e.g. Lachish et al., 2012; Miller et al., 2012).

Imperfect pathogen detection has been widely acknowledged in both veterinary and medical fields and is likely present in most sampling and diagnostic methods used by disease ecologists. Veterinary and medical fields have long used statistical tools to adjust pathogen prevalence estimates by correcting for the accuracy of diagnostic tests (reviewed in Enoe, Georgiadis, & Johnson, 2000; Greiner & Gardner, 2000; Toft, Jørgensen, & Højsgaard, 2005). However, the stringent assumptions and requirements of these statistical tools make them impractical for disease ecologists. For example, most methods used in veterinary and medical fields involve determining the accuracy of a diagnostic test by comparing it to an independent reference test (e.g. Drewe, Dean, Michel, & Pearce, 2009; Greiner & Gardner, 2000). In the realm of disease ecology, most disease diagnostics are obtained from a single test or from visual inspections when no diagnostic tools are available (e.g. facial tumour disease of Tasmanian devils, *Sarcophilus harrisii*; Lachish, Jones, & McCallum, 2007).

In the case of pathogen presence, uncertainty is related to the specificity (i.e. the probability an uninfected individual is correctly classified as uninfected) and sensitivity (i.e. the probability an infected host is correctly classified as infected) of the sampling and diagnostic methods. Typically, specificity is assumed maximised when strict protocols are used in the field and lab to decrease the odds of contaminating samples that lead to false positives. False negatives, alternatively, occur during a survey event when the pathogen is present but is not detected (e.g. Colvin, Peterson, Kent, & Schreck, 2015; Thompson, 2007). Sensitivity, therefore, is the product of two processes: (1) sampling methods (e.g. blood, swab, histology sample, etc.) and (2) laboratory diagnostic testing (e.g. qPCR, ELISA, etc.). For example, the causative agent of whirling disease, *Myxobolus cerebralis*, infects the brain of a fish, and infections can be missed when an uninfected area of the brain is examined (Thompson, 2007). In this case, it is also likely that imperfect pathogen detection is related to pathogen infection intensity (e.g. Valkiunas et al., 2008), where low-level infections are more likely missed than high-level infections. Few field studies, however, consider false negative error rates of sampling methods, and even fewer directly estimate them (e.g. Colvin et al., 2015; Thompson, 2007).

Thus far, the primary focus of disease ecologists investigating false negative error rates of pathogens has occurred with respect to laboratory diagnostic tests. For example, several studies have investigated how the sensitivity of quantitative PCR depends on host infection intensity; as host infection intensity increases, the probability of detecting the pathogen also increases (e.g. Gómez Díaz, Doherty, Duneau, & McCoy, 2010; Lachish et al., 2012; Miller et al., 2012). This pattern has been detected across several disease systems using different diagnostic tests, including: qPCR to detect the causative agent of malaria, *Plasmodium* sp., in birds (Knowles et al., 2011; Lachish et al., 2012); qPCR to detect *Batrachochytrium dendrobatidis* on amphibian skin (Miller et al., 2012); γ interferon and ELISA tests to detect the causative agent of tuberculosis, *Mycobacterium bovis*, in cattle (Ritacco et al., 1991); and qPCR to detect the causative agent of Lyme disease, *Borrelia* species complex, in *Ixodes uriae* ticks (Gómez Díaz et al., 2010). Cumulatively, this evidence strongly suggests that host infection intensity affects the probability of detecting the pathogen using several different diagnostic tests, but it remains unclear if host infection intensity affects the probability of detecting the pathogen during sampling.

As a motivating example, we focus on the emerging infectious fungal pathogen *Batrachochytrium dendrobatidis* (hereafter *Bd*; Longcore, Pessier, & Nichols, 1999), the causative agent of chytridiomycosis in amphibians. *Bd* is one of the greatest threats to amphibian biodiversity; it has been detected on over 700 amphibian species; and it has been found on every continent where amphibians occur (Cheng, Rovito, Wake, & Vredenburg, 2011; Fisher, Garner, & Walker, 2009; Lips et al., 2006; Olson et al., 2013; Wake & Vredenburg, 2008). To date, the most sensitive sampling and diagnostic methods to test for the presence of *Bd* are non-invasive skin swabs and qPCR (Kriger, Hero, & Ashton, 2006). While it has been shown that, like most other diagnostic tests, qPCR sensitivity to *Bd* is <1 and correlates with host infection intensity (e.g. Miller et al., 2012), it remains unclear if host infection intensity also impacts *Bd* sampling sensitivity of non-invasive skin swabs (i.e., replication frequency, number of swab strokes, pressure of swab, etc.). Abundance-induced detection heterogeneity is

well-known to affect the estimation of occurrence and abundance (e.g. Royle & Nichols, 2003), so it is expected that if two swabs were collected in sequence from the same amphibian, the likelihood that both swabs would detect the pathogen and quantify the same infection intensity should be lower at low-level infection intensities (e.g. Lachish et al., 2012; Miller et al., 2012).

In this paper, our objectives were two-fold. First, we quantified false negative error rates from imperfect host sampling (via non-invasive skin swabbing) of *Bd* in an amphibian community in El Copé, Panama. To do this, we swabbed amphibians twice in sequence, and we used a recently developed hierarchical Bayesian estimator formulated by Miller et al. (2012), originally used to examine qPCR false negative rates of *Bd* on amphibians. We expected that as host infection intensity increased, the probability of detecting *Bd* on a skin swab would increase, similar to the relationship between *Bd* infection intensity and qPCR *Bd* detection probability (Miller et al., 2012). We also assessed the variation in *Bd* prevalence and infection intensity between habitats (stream vs. trail) and seasons (wet vs. dry) because previous studies have shown that these variables explain differences in host disease susceptibility (e.g. Brem & Lips, 2008; Kriger & Hero, 2006). Second, we developed a novel hierarchical Bayesian model that simultaneously accounted for imperfect pathogen detection from both field sampling and diagnostic tests. We simulated and analysed data under a variety of sampling design scenarios to quantify the magnitude of parameter bias of pathogen prevalence and infection intensity estimates.

Our modelling approach and results provide tools for disease ecologists to refine and optimise pathogen-sampling procedures and reduce the bias of parameter estimates, which will improve inference and the application of epidemiological models to understand and forecast host–pathogen dynamics. We provide an appendix with R code to facilitate the application of these methods. However, we highlight that the biases introduced in estimating parameters of interest and methodological recommendations is highly dependent on details of the study system and objectives. Our driving motivation in the development and applications of this method is to understand how imperfect pathogen detection from samples and diagnostic tests contribute to biases in population-level inferences, which should guide the efficient allocation of resources in epidemiological studies.

2 | MATERIALS AND METHODS

2.1 | Field surveys

We sampled four-200 m stream and three-400 m trail transects in Parque Nacional G. D. Omar Torríjos Herrera, Coclé Province, El Copé, Panama (8°40′ N, 80°37′17″ W; Lips, Reeve, & Witters, 2003) during two wet seasons (2012, 2013) and one dry season (2013). The park spans elevations between 500 and 1,000 m and is located on the Continental Divide. This site experiences both dry (December–April) and wet (May–November) seasons.

We surveyed each transect six to eight times during each season. Field teams of two to three people conducted nocturnal visual encounter surveys by slowly walking each transect and using visual and audio cues to locate amphibians within two metres of the stream bank or trail. Upon capture, we swabbed the abdomen and each limb five times (total = 30 strokes) per animal following the swabbing protocol by Hyatt et al. (2007) using a sterile cotton-tipped swab (Dry Swab MW113, Medical Wire). We collected at least one swab per individual captured, and a subset of individuals was swabbed twice in sequence from the same location on the animal's body (see Table 1a in Appendix S1). We used a fresh pair of latex powder-free gloves when handling each individual. We stored all swabs in individually capped 2 ml tubes with 30 μl of 70% ethanol. Because we did not uniquely mark all individuals that we captured and swabbed, it is possible that we repeatedly swabbed the same individual within a season, making some samples pseudo-replicates. Pseudo-replicates, in this case, will decrease the variability around our reported naïve and adjusted estimates of pathogen prevalence and infection intensity. All individuals were released at the original point of capture. We include *Bd* infection intensity data from all amphibian species captured (see Table 1a in Appendix S1) without discriminating among species in the model because <10 individuals per species were captured for *c*. 56% of the 39 total species detected.

2.2 | Molecular analysis

We used PrepMan Ultra® for DNA preparation of swabs tested for *Bd*. We tested swabs for *Bd* in singlicate, using Taqman qPCR (Boyle, Boyle, Olsen, Morgan, & Hyatt, 2004; Hyatt et al., 2007) running 50 cycles. We ran each plate with JEL 423 standards of 0.1, 1, 10, 100, and 1,000 *Bd* zoospore genomic equivalents (ZGE) to determine *Bd* presence and infection intensity. Isolate JEL 423 was originally isolated at El Copé, Panama during the epizootic of 2004. We categorised individuals as *Bd*-positive when infection intensity was greater than zero (Briggs, Knapp, & Vredenburg, 2010). The qPCR assay consistently detects very small *Bd* infections (0–1 ZGE), likely representing very low levels of infection. To ensure that false positives were negligible, we included multiple negative controls in each qPCR plate.

2.3 | Sampling detection-adjusted model

We used a slightly modified version of the hierarchical Bayesian estimator developed by Miller et al. (2012) to account for heterogeneity in pathogen detection probability due to host infection intensity. We assumed that there was no error associated with the diagnostic test. Instead, we focus on the sensitivity of the swabbing procedure. We did not use the multi-season formulation of the occupancy model because we did not track the individuals that were swabbed across seasons. Alternatively, we assumed that infection intensity solely depended on habitat and season and not an individual's previous infection history.

We modelled the true, but unobservable, *Bd* infection state (z_i) on the *i*th individual as:

$$z_i \sim \text{Bernoulli}(\Psi_{\text{habitat}_i,\text{season}_i}) \quad (1)$$

where an individual was either infected ($z_i = 1$) or not ($z_i = 0$), and Ψ, the probability of the *i*th individual being infected, depends on habitat type (trail vs. stream) and season (dry vs. wet). Note that the parameter Ψ gives rise to the estimate of pathogen prevalence, i.e., the proportion of hosts infected, which is the outcome from the repeated Bernoulli process.

We modelled the observed *Bd* infection on the *i*th individual and the *j*th swab, y_{ij}, as:

$$y_{ij} \sim \text{Bernoulli}(p_i \times z_i), \quad (2)$$

where an infection was either detected ($y_{ij} = 1$) or not ($y_{ij} = 0$). We modelled the detection probability, p_i, as a function of true, but unobservable, *Bd* infection intensity, x_i, on the *i*th individual as:

$$\text{logit}(p_i) = \alpha + \beta \times x_i \quad (3)$$

where α is the log odds of pathogen detection when infection intensity x_i is zero, and β is the scaling coefficient representing how detection log odds changes with respect to host infection intensity. We modelled the true log *Bd* infection intensity, x_i, on the *i*th host as:

$$\log(x_i) \sim \text{normal}(\mu_i, \sigma^2) \quad (4)$$

where the true mean log infection intensity, μ_i, was a function of habitat, season, and true *Bd* infection state, z_i, for the *i*th host:

$$\mu_i = \log(0.001 + \omega_{\text{habitat}_i,season_i} \times z_i) \quad (5)$$

Above, σ^2 represents the standard deviation in *Bd* infection intensities across the host population.

Lastly, we modelled the observed infection intensity, w_{ij}, on the *i*th individual and the *j*th swab as:

$$\log(w_{ij}) \sim \text{normal}(\log(x_i + 0.001), \sigma^2_{\text{error}}) \quad (6)$$

In this case, σ^2_{error} represents the measurement error of the estimates for *Bd* infection intensity produced by the non-invasive swabbing technique.

2.4 | Unadjusted model

To estimate the parameter bias caused by pathogen non-detection, we fit the same model outlined above after removing the detection probability and measurement error portions of the model (Equations 2 and 6). We modified our data by collapsing the host by swab matrix in two key ways: (1) if *Bd* was detected on any swab collected for an individual then that individual was considered infected, and (2) we averaged the infection intensities across all swabs for each host that was considered infected.

2.5 | Model fit

We fit all models using Bayesian methods and estimated the posterior distributions for all parameters and latent states using Markov chain Monte Carlo (MCMC) implemented in JAGS 4.0.0 with the JAGSUI package (Kellner, 2015) in the R environment (R Core Team, 2015). We used vague priors (i.e., normal(0, 0.01) or normal(0, 0.368); Lunn, Jackson, Best, Thomas, & Spiegelhalter, 2012) for all parameters. We computed three chains for each random variable with diffuse initial values. After a burn-in of 10,000 iterations, we accumulated 40,000 samples from each chain, keeping every 50th sample. We assessed convergence by visually inspecting trace plots and using the diagnostics of Gelman (Brooks & Gelman, 1998). We used a posterior predictive check (hereafter Bayesian *p*-value) to compare the observed data to simulated datasets generated from the parameter estimates at each step in the MCMC algorithm. We confirmed that the Bayesian *p*-value, defined as the probability that the simulated data were more extreme than the observed data, was indicative of a good model fit (e.g. Gelman et al., 2013; Kéry & Schaub, 2012). Our observed data fit both the sampling detection-adjusted and unadjusted models well (see Figure 1a, Bayesian *p-value* = .79; Figure 2a, Bayesian *p-value* = .41 in Appendix S1).

To quantify effects, we calculated the differences between parameters of interest at each MCMC iteration following Ruiz-Gutiérrez, Zipkin, and Dhondt (2010). We computed the proportion of iterations where one parameter was greater than the other, which is directly interpreted as the probability (Pr) that one parameter is greater than the other. We considered effects with large credible intervals to be either unimportant to the process being modelled, or to have been estimated too imprecisely to draw conclusive inference.

2.6 | *Bd*-specific methodological guidance

Given that *Bd* infection may be overlooked on an infected host, we followed Kery (2002) and calculated the probability of detecting *Bd*, P^*, on *n* identical and independent swabs or qPCR runs using the binomial argument:

$$P^* = 1 - (1 - p)^n, \quad (7)$$

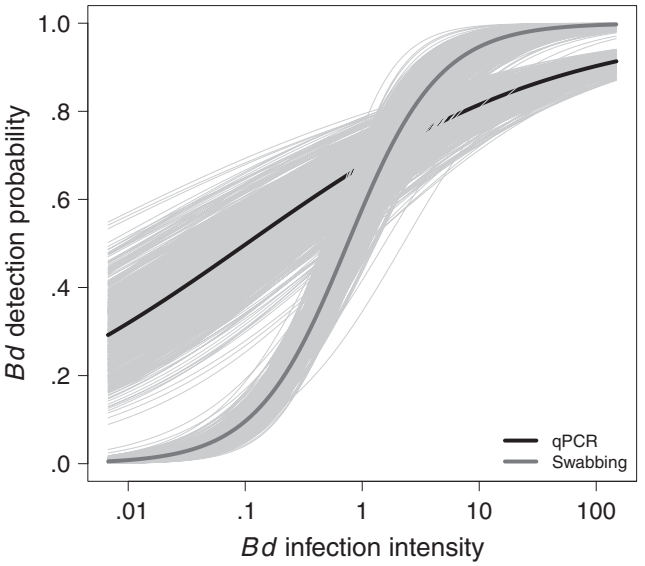

FIGURE 1 The relationship between *Batrachochytrium dendrobatidis* (*Bd*) detection probability and host infection intensity caused by laboratory (black line; qPCR error; Miller et al., 2012) and swabbing (dark grey line) methods. This graph indicates that as host infection intensity increases, pathogen detection probability also increases. The dark lines are mean posterior distribution estimates, and light grey lines represent the 95% credible interval around the mean

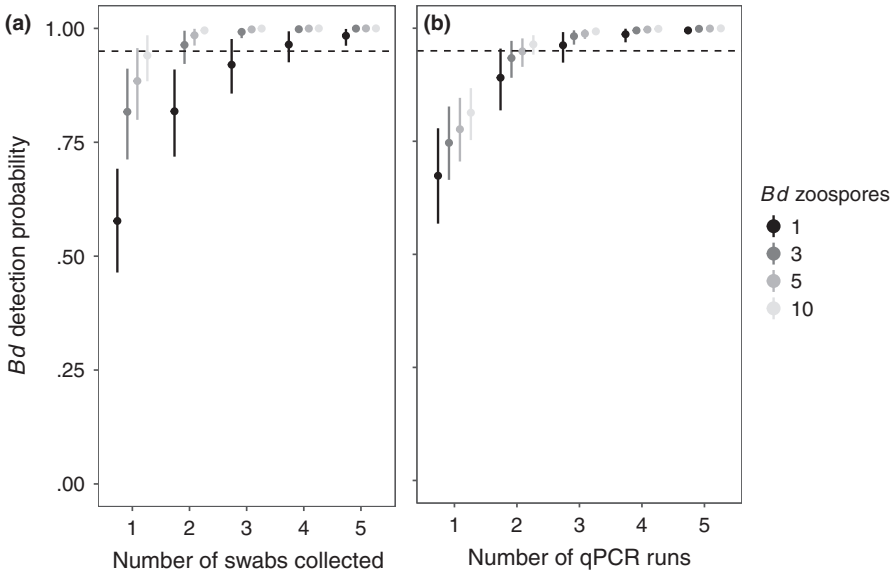

FIGURE 2 Given that pathogen infection may be overlooked on an infected individual, we calculated the probability of detecting *Batrachochytrium dendrobatidis* (*Bd*) on *n* identical and independent (a) swabs or (b) qPCR runs, using a binomial argument, as $P^* = 1 - (1 - p)^n$ (Kery, 2002). We model the probability of detecting 1 ZGE, 3 ZGE, 5 ZGE, and 10 ZGE. The dashed line indicates 95% certainty of detecting the pathogen when it is present

where *p* was either obtained from the sampling detection-adjusted model outlined above (i.e., imperfect *Bd* detection from swabbing) or from Miller et al. (2012; i.e., imperfect *Bd* detection from qPCR). We modelled the probability of detecting 1 ZGE, 3 ZGE, 5 ZGE, and 10 ZGE. Our recommendations were based on the minimum number of swabs and qPCR runs to be 95% certain that if *Bd* were present then it would be detected.

2.7 | Sampling and diagnostic detection-adjusted model

To provide general methodological guidance, we developed a novel Bayesian hierarchical model that simultaneously accounts for imperfect sampling and diagnostic detection of the pathogen (see Appendix S2 for the full model outline and R code). We performed a simulation study to explore the ability of this hierarchical model to estimate pathogen prevalence and infection intensity under scenarios of imperfect sampling and diagnostic testing for a pathogen. We set average infection intensity as either low ($\mu = 2$) or high ($\mu = 20$) and the probability of pathogen infection as either low ($\Psi = 0.20$) or high ($\Psi = 0.80$). We assumed low population infection intensity standard deviation ($\sigma^2 = 1$), infection intensity measurement error in sampling ($\sigma^2_{error} = 1$), and infection intensity measurement error of the diagnostic test ($\sigma^2_{diagnostic} = 1$). We set the log odds scaling coefficient of pathogen detection for sampling and diagnostic testing as either low (-1.38) or high (1.38). We did not consider the case where detection probability for either pathogen sampling or diagnostic testing varied with respect to true infection intensity (i.e., changing the slopes of the relationships). We explored how varying the number of samples collected per host (i.e., 1, 2, 3, 4) and the number of diagnostic runs per sample (i.e., 1, 2, 3, 4) affected the bias in estimated parameters of interest (i.e. pathogen prevalence [Ψ] and average infection intensity [μ]). This resulted in 256 parameter combinations (see Table 2a in Appendix S1). For each parameter combination, we simulated 50 datasets of 500 individuals each before fitting the model using the same methods outlined above. To quantify the magnitude of parameter bias under each scenario, we calculated the root mean squared error between the posterior mean and the actual parameter value.

TABLE 1 Summary of the posterior distributions from the sampling detection-adjusted model. All parameters were back transformed to their original scale, except detection probability (logit scale) and error estimates

Definition	*Mean*	95% credible interval	
ω			
Stream dry	0.13	0.04	0.43
Trail dry	0.40	0.11	1.39
Stream wet	0.13	0.02	1.12
Trail wet	0.16	0.07	0.39
Process error (σ^2)	3.17	2.66	3.63
Measurement error (σ^2_{error})	1.03	0.81	1.29
Ψ			
Stream dry	0.86	0.57	0.98
Trail dry	0.67	0.46	0.91
Stream wet	0.64	0.31	0.95
Trail wet	0.92	0.73	0.99
Detection probability			
α	0.22	−0.23	0.68
β	0.93	0.67	1.20

3 | RESULTS

3.1 | Field summary

We captured and swabbed 865 individuals of 39 species at least once (see Table 1a in Appendix S1). We collected 148 and 99 swabs on streams and trails, respectively, during the dry season, and 288 and

TABLE 2 Summary of the posterior distribution from the unadjusted model. All parameters were back transformed to their original scale

Definition	*Mean*	95% credible interval	
ω			
Stream dry	0.71	0.38	1.34
Trail dry	1.77	1.18	2.65
Stream wet	1.33	0.50	3.36
Trail wet	1.32	0.91	1.96
Ψ			
Stream dry	0.30	0.23	0.37
Trail dry	0.28	0.23	0.34
Stream wet	0.21	0.14	0.30
Trail wet	0.33	0.28	0.39

302 swabs on streams and trails, respectively, during the two wet seasons. Of the 865 individuals, we double swabbed 205 individuals, where 102 were double swabbed during the dry season and 103 were double swabbed during the two wet seasons. Of these 205 double swabbed individuals, we detected *Bd* DNA on only one swab from 51 individuals, and on both swabs from 25 individuals.

3.2 | Sampling detection-adjusted model

Under the sampling detection-adjusted model, the probability of being infected with *Bd* did not differ between streams and trails during the wet season ($Pr(\Psi_{stream,wet} > \Psi_{trail,wet}) = 0.10$) nor the dry season ($Pr(\Psi_{stream,dry} > \Psi_{trail,dry}) = 0.93$; Table 1). When comparing the probability of being infected with *Bd* in particular habitat types between seasons, the probability of being infected with *Bd* did not differ between wet and dry seasons for streams ($Pr(\Psi_{stream,wet} > \Psi_{stream,dry}) = 0.38$), but it did differ for trails ($Pr(\Psi_{trail,wet} > \Psi_{trail,dry}) = 0.98$).

Average infection intensity did not differ between streams and trails during the wet season ($Pr(\omega_{stream,wet} > \omega_{trail,wet}) = 0.85$) nor the dry season ($Pr(\omega_{stream,dry} > \omega_{trail,dry}) = 0.14$). When comparing average infection intensity in particular habitat types between seasons, average infection intensity differed between wet and dry seasons for streams ($Pr(\omega_{stream,wet} > \omega_{stream,dry}) = 0.99$) but not for trails ($Pr(\omega_{trail,wet} > \omega_{trail,dry}) = 0.49$).

Bd detection probability increased as host infection intensity increased (Figure 1). *Bd* detection probability was *c*. 99.99% at an infection intensity of 10 ZGE (Figure 1).

3.3 | Unadjusted model

Contrary to the sampling detection-adjusted model, the unadjusted model revealed that the probability of being infected with *Bd* differed between streams and trails only during the wet ($Pr(\Psi_{stream,wet} > \Psi_{trail,wet}) = 0.03$) but not during the dry season ($Pr(\Psi_{stream,dry} > \Psi_{trail,dry}) = 0.84$; Table 2). In contrast, when comparing habitat types between seasons, the probability of being infected with *Bd* did not differ between wet and dry seasons for streams ($Pr(\Psi_{stream,wet} > \Psi_{stream,dry}) = 0.12$) nor trails ($Pr(\Psi_{trail,wet} > \Psi_{trail,dry}) = 0.92$), which is similar to the sampling detection-adjusted model.

Again, similar to the sampling detection-adjusted model, average infection intensity from the unadjusted model did not differ between streams and trails during the wet season ($Pr(\omega_{stream,wet} > \omega_{trail,wet}) = 0.91$) nor the dry season ($Pr(\omega_{stream,dry} > \omega_{trail,dry}) = 0.49$). In contrast to the sampling detection-adjusted model, average infection intensity did not differ between wet and dry seasons for streams ($Pr(\omega_{stream,wet} > \omega_{stream,dry}) = 0.93$) nor trails ($Pr(\omega_{trail,wet} > \omega_{trail,dry}) = 0.60$).

3.4 | Sampling detection-adjusted model vs. unadjusted model

All four of the Ψ parameters estimates, quantifying the probability of being infected with *Bd*, from the unadjusted model were lower than the parameter estimates from the sampling detection-adjusted model (all $Pr(\Psi_{adjusted} > \Psi_{unadjusted}) < 0.05$). Likewise, all of the parameter estimates for average infection intensity, ω, from the unadjusted model were higher than the sampling detection-adjusted model (all $Pr(\omega_{adjusted} > \omega_{unadjusted}) > 0.95$).

3.5 | *Bd*-specific methodological guidance

To be 95% certain that 1 ZGE is present on a host and that it is detected using non-invasive skin swabs, at least four swabs need to be collected. While, to be 95% certain that 3 or 5 ZGE are present and detected, at least two swabs need to be collected (Figure 2). On the contrary, non-invasive skin swabs can detect 10 ZGE with greater than 95% certainty using only a single skin swab. To be 95% certain that 1, 3, 5, or 10 ZGE are present on a host and detected using qPCR, at least two qPCR runs need be performed per sample collected (Figure 2).

3.6 | Sampling and diagnostic detection-adjusted model

The estimated probability of pathogen infection, Ψ, was less biased and more precise when average infection intensity, μ, was high and when both contributors to pathogen detection probability—sampling methods and laboratory diagnostic testing—were high (Figure 3; see Figure 3a in Appendix S1). In general, the root mean squared error of pathogen prevalence decreased more rapidly when the number of samples increased rather than the number of diagnostic tests in most scenarios (Figure 3).

Similarly, estimated average infection intensity, μ, was less biased and more precise when the probability of pathogen infection, Ψ, was high and at high values of pathogen detection probability (Figure 4; see Figure 4a in Appendix S1). In general, the root mean squared error of the estimated infection intensity was similarly impacted if either the number of samples collected or the number of diagnostic runs increased (Figure 4).

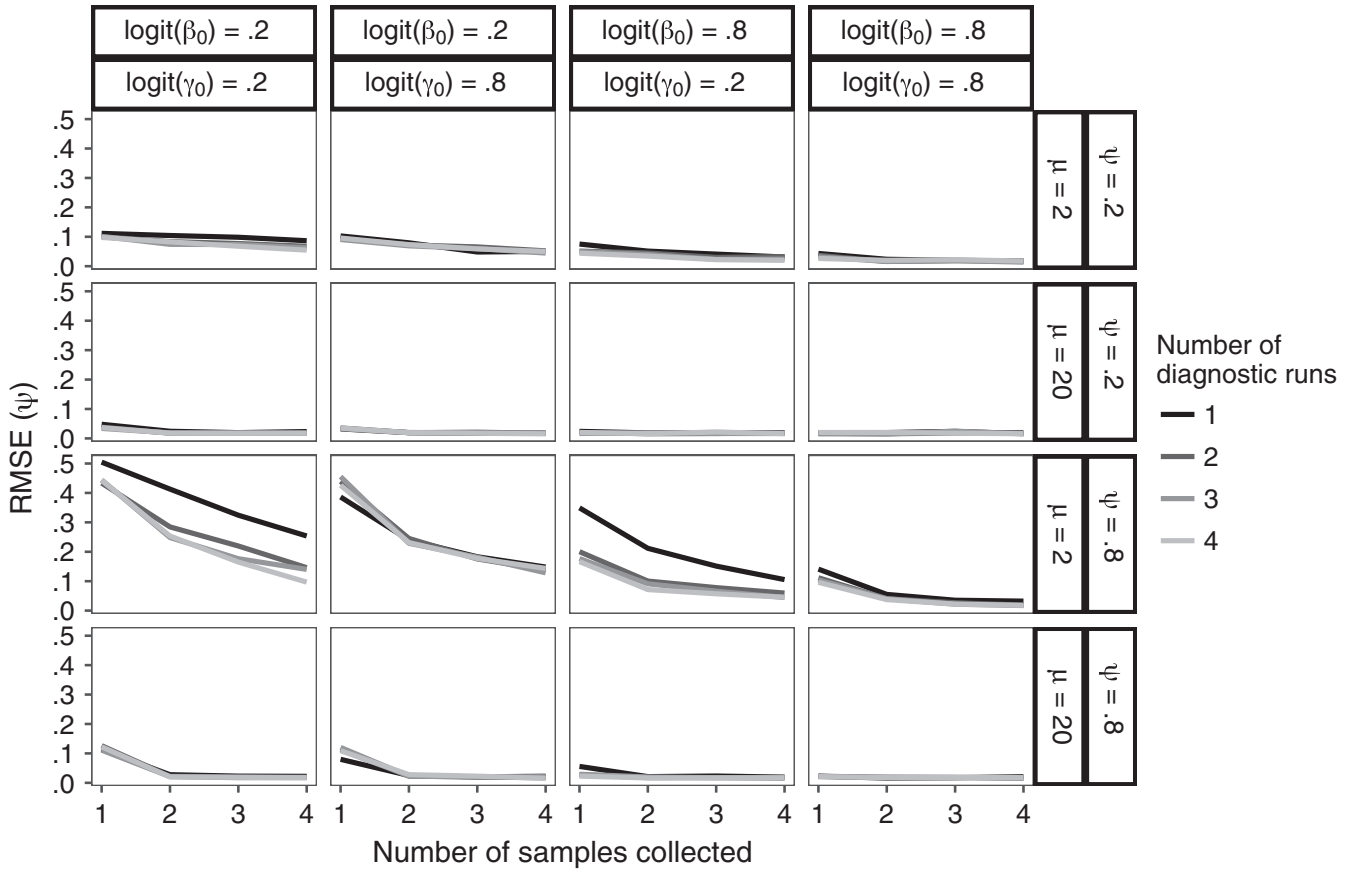

FIGURE 3 Root mean square error of estimated pathogen prevalence (Ψ) over different scenarios of known high and low pathogen prevalence (Ψ), average infection intensity (μ), pathogen log odds of detection by sampling method (β_0), and pathogen log odds of detection by laboratory diagnostics (γ_0) as the number of samples (1–4) and number of diagnostic runs (1–4) vary

4 | DISCUSSION

4.1 | Implications for wildlife disease ecology

Our results demonstrate that non-invasive skin swabbing imperfectly detects *Bd* and that Bayesian hierarchical models can adjust prevalence and average infection intensity for low-level pathogen infections that are missed. The bias caused by pathogen non-detection can affect disease inference, especially in regions where hosts harbour low-level *Bd* infections, such as in enzootic populations.

Imperfect pathogen detection threatens the success of disease monitoring programs intended to prevent pathogen invasion (e.g. Langwig et al., 2015). For example, the recent emergence of *Batrachochytrium salamandrivorans* (hereafter *Bsal*; Martel et al., 2013), the only known sister taxa of the amphibian-killing fungus, *Bd* (Longcore et al., 1999), threatens salamander biodiversity worldwide and is lethal to some of the New World salamandrid species (genera: *Taricha* and *Notophthalmus*; Martel et al., 2014). The United States has taken precautionary measures to prevent the arrival of *Bsal* into its borders by restricting the movement of salamanders under the Lacey Act (18 U.S.C. 42). The techniques used to sample and diagnose *Bsal* are similar to those used to test for *Bd*, such as non-invasive skin swabs tested by qPCR (Hyatt et al., 2007; Martel et al., 2013). At some point, a second precautionary step would be to require that salamanders obtain health certificates to move across borders. But, given the results of this study, there is a chance that low-level *Bsal* infections will be missed. If *Bsal* detection probability is similar to *Bd*, then we expect that *Bsal* infections less than 5 ZGE will likely be missed up with a certainty of 95% by a single non-invasive skin swab; similarly, qPCR will detect infections less than 10 ZGE approximately 81% of the time if only one qPCR run were performed. This is especially concerning when importing salamanders from Eastern Asia, where salamanders typically have *Bsal* infection intensities less than 30 zoospores (Martel et al., 2014) and *Bsal* prevalence is low (i.e., <10%; Laking, Ngo, Pasmans, Martel, & Nguyen, 2017; Martel et al., 2014).

In El Copé, Panama, where *Bd* is now enzootic, most individuals were infected (average *Bd* prevalence *c*. 64%–92%) and carried low-level infections (<10 ZGE), which is similar to other regions in the Americas (James et al., 2015). In the 1990s, as *Bd* spread worldwide, many amphibian populations experienced mass mortality events and population declines (e.g. Berger, Hyatt, Speare, & Longcore, 2005; Lips et al., 2006; Muths, Corn, Pessier, & Green, 2003; Vredenburg, Knapp, Tunstall, & Briggs, 2010). In many of these areas today, amphibians persist with enzootic *Bd* infections, and disease ecologists are interested in explaining the ecological patterns of infection and

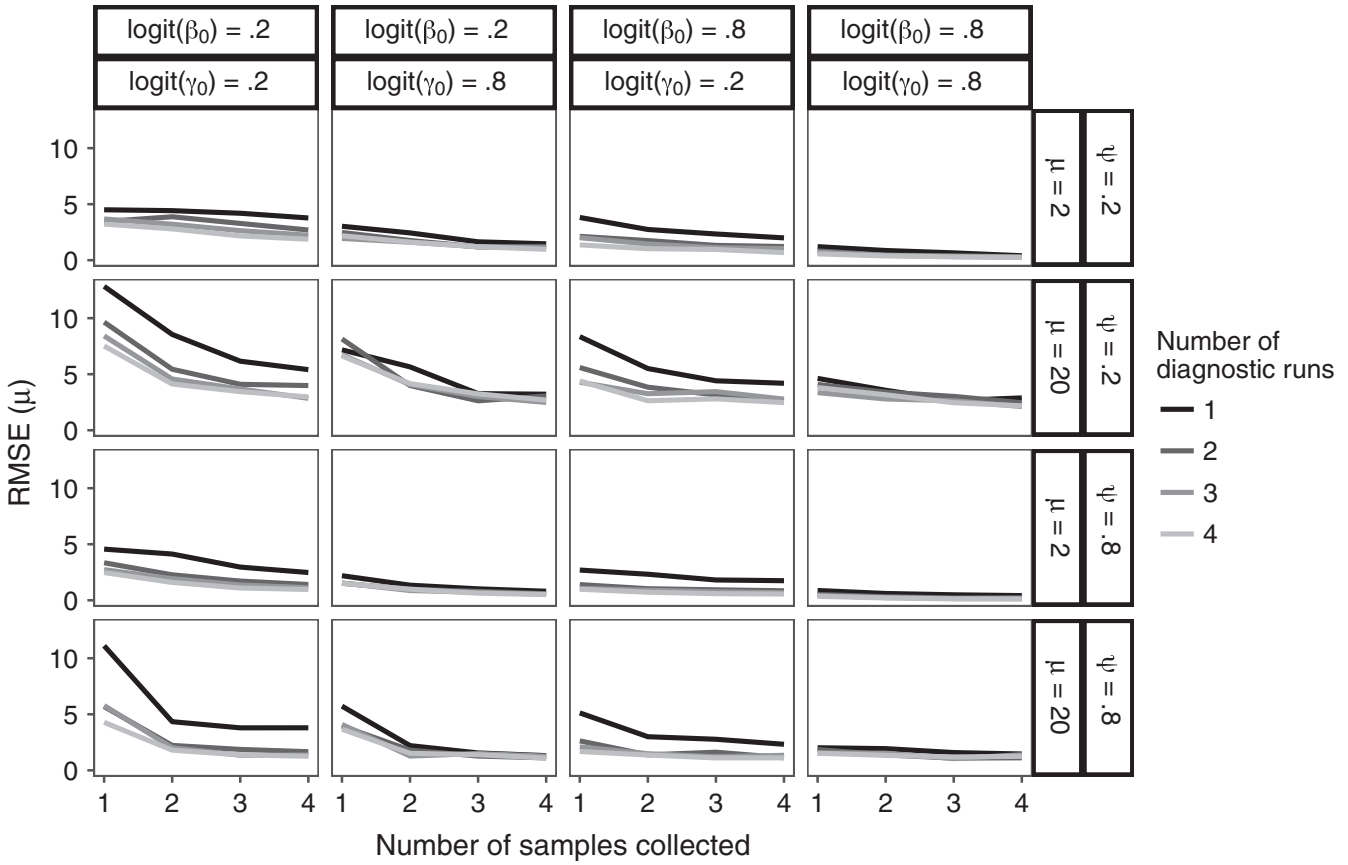

FIGURE 4 Root mean square error of estimated average infection intensity (μ) over different scenarios of known high and low pathogen prevalence (Ψ), average infection intensity (μ), pathogen log odds of detection by sampling methods (β_0), and pathogen log odds of detection by laboratory diagnostics (γ_0) as the number of field samples (1–4) and number of diagnostic runs (1–4) vary

host–pathogen coexistence. If disease ecologists do not correct for imperfect pathogen detection, then estimates of disease will be biased, and in some cases, covariates that affect the sampling process may end up in the ecological process model (e.g. Kéry & Schaub, 2012). In this study, our conclusions on *Bd* probability of infection and infection intensity with respect to habitat and seasons were predominately similar regardless of whether we accounted for imperfect pathogen detection, which is not surprising given the "noisy" estimates for these covariates. However, the parameter estimates from the sampling detection-adjusted and unadjusted model differed significantly with respect to precision and bias, which emphasises the importance of accounting for imperfect pathogen sampling. If overlooked, even small frequencies of false negatives can lead to inaccurate inference and biased conclusions.

4.2 | *Bd*-specific methodological guidance

In the case of *Bd*, we found that the greatest pathogen detection bias was caused when host infection intensity was low, as expected. We highlight that pathogen detection probability is lower than reported from only the double swab data because total imperfect pathogen detection depends on both pathogen detection probability of the laboratory diagnostic tests (i.e., qPCR; Lachish et al., 2012; Miller et al., 2012) and sampling methods (i.e., swabbing; e.g. Thompson, 2007; Figure 1). These results indicate that replication of samples in both sampling and laboratory methods are critical to minimise observational uncertainty, especially when pathogen prevalence and infection intensity are expected to be low. This is the case in both enzootics and in the invasion phase of an epizootic (Langwig et al., 2015).

We recognise the increase cost and effort needed to analyse more swab samples in replicate; therefore, we suggest collecting replicate swabs when possible because if the results from the first swab set shows few pathogen detections, low pathogen prevalence, and low host infection intensity, it may be worth analysing the second set to calculate false negative error rates.

4.3 | General methodological guidance

Applying the sampling and diagnostic detection-adjusted model, we find that there are trade-offs, for a fixed effort, in precision and accuracy of pathogen prevalence and average infection intensity estimates. Although our simulation study provides general methodological guidance under different sampling scenarios, these results must be considered in combination with common sense and expert knowledge of the study system. For example, researchers must consider the cost and time constraints of collecting multiple samples per individual and running multiple diagnostic tests per sample, as well as the trade-offs between sampling breadth and accuracy. Given that the magnitude of bias depends on the study

system, type of infection, and false negative error rates, we advocate that the most cautious approach is to assume all errors are possible and to accommodate them by adjusting sampling designs. Using the R code provided in Appendix S2 as a foundation, it may be worthwhile to simulate study-specific scenarios to understand the trade-offs between efficiency and robustness of particular study designs.

The models we present here can be applied to designing studies and analysing data for other emerging infectious fungal diseases, such as white-nose syndrome (Langwig et al., 2015), snake fungal disease (Tetzlaff, Allender, Ravesi, & Smith, 2015), and the salamander fungus (*Bsal*; Martel et al., 2014). Similar to detecting *Bd* on amphibian skin, each of these study systems is typified by the collection of non-invasive skin samples and using qPCR analysis to test for pathogen presence and infection intensity, subjecting them to similar kinds of detection errors as the *Bd*-amphibian system. Should these infectious diseases spread from their initial distributions, it will be critical to be able to compare disease status and dynamics across studies with respect to biotic and abiotic covariates. This will require unbiased and precise estimates of key epidemiological parameters, such as pathogen prevalence and infection intensity. By accounting for detection errors arising from sampling and diagnostic tests, we can more readily compare disease inference among systems and species.

ACKNOWLEDGEMENTS

We thank N. Angeli, A. Cunha, C. Johnston, E. Kabay, C. Muletz, J. Ray, and T. Tunstall for help in the field. We thank L. Browne and S. Saunders for help on previous drafts. We also thank three anonymous reviewers and the editor for greatly improving the quality of the manuscript. Animal use and collection was approved by the University of Maryland IACUC (Protocol no.: R-12-05), permits from the Autoridad Nacional Del Ambiente (no.: SE/AH-3-12), and the Smithsonian Tropical Research Institute. This work was funded by a National Science Foundation Graduate Research Fellowship to G.V.D. and grants from the University of Maryland Latin-American Studies office and the National Science Foundation (DEB no. 1120161) to K.R.L. and K.R.Z. This is contribution number 602 of the Amphibian Research and Monitoring and Initiative (ARMI). There is no conflict of interest among authors.

AUTHORS' CONTRIBUTIONS

G.V.D. helped in designing the study, collected and analysed samples in the laboratory, formulated the models, and wrote the first draft of the paper. K.R.L. helped design the study and collect samples. A.V.L. and K.R.Z. helped analyze the samples in the laboratory. C.C.C. helped formulate the models. E.H.C.G. helped design the study and formulate the models. All authors contributed substantially to manuscript revisions.

DATA ACCESSIBILITY

The data used in the analysis for the main text can be found online at the Dryad Digital Repository https://doi.org/10.5061/dryad.p1006 (DiRenzo et al., 2017). Simulation data can be generated using code in Appendix S2, and all the R scripts used to analyse the data and create figures can be found on the Github repository https://github.com/Grace89/ImperfectPathogenDetection_MEES. This repository is citable using: https://doi.org/10.5281/zenodo.840132.

ORCID

Graziella V. DiRenzo http://orcid.org/0000-0001-5264-4762
Evan H. Campbell Grant http://orcid.org/0000-0003-4401-6496

REFERENCES

Berger, L., Hyatt, A. D., Speare, R., & Longcore, J. E. (2005). Life cycle stages of the amphibian chytrid *Batrachochytrium dendrobatidis*. *Diseases of Aquatic Organisms*, *68*, 51–63.

Boyle, D. G., Boyle, D. B., Olsen, V., Morgan, J. A. T., & Hyatt, A. D. (2004). Rapid quantitative detection of chytridiomycosis. *Diseases of Aquatic Organisms*, *60*, 141–148.

Brem, F., & Lips, K. (2008). *Batrachochytrium dendrobatidis* infection patterns among Panamanian amphibian species, habitats and elevations during epizootic and enzootic stages. *Diseases of Aquatic Organisms*, *81*, 189–202.

Briggs, C. J., Knapp, R. A., & Vredenburg, V. T. (2010). Enzootic and epizootic dynamics of the chytrid fungal pathogen of amphibians. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 9695–9700.

Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7*, 434–455.

Cheng, T. L., Rovito, S. M., Wake, D. B., & Vredenburg, V. T. (2011). Coincident mass extirpation of neotropical amphibians with the emergence of the infectious fungal pathogen *Batrachochytrium dendrobatidis*. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 9502–9507.

Colvin, M. E., Peterson, J. T., Kent, M. L., & Schreck, C. B. (2015). Occupancy modeling for improved accuracy and understanding of pathogen prevalence and dynamics. *PLoS ONE*, *10*, e0116605.

Cooch, E. G., Conn, P. B., Ellner, S. P., Dobson, A. P., & Pollock, K. H. (2011). Disease dynamics in wild populations: Modeling and estimation: A review. *Journal of Ornithology*, *152*, 485–509.

DiRenzo, G. V., Campbell Grant, E. H., Longo, A. V., Che-Castaldo, C., Zamudio, K. R., & Lips, K. R. (2017). Data from: Imperfect pathogen detection from non-invasive skin swabs biases disease inference. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.p1006

Drewe, J. A., Dean, G. S., Michel, A. L., & Pearce, G. P. (2009). Accuracy of three diagnostic tests for determining *Mycobacterium bovis* infection status in live-sampled wild meerkats (*Suricata suricatta*). *Journal of Veterinary Diagnostic Investigation*, *21*, 31–39.

Enoe, C., Georgiadis, M. P., & Johnson, W. O. (2000). Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease statue is unknown. *Preventive Veterinary Medicine*, *45*, 61–81.

Fisher, M. C., Garner, T. W. J., & Walker, S. F. (2009). Global emergence of *Batrachochytrium dendrobatidis* and amphibian chytridiomycosis in space, time, and host. *Annual Review of Microbiology*, *63*, 291–310.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. Boca Raton, FL: CRC Press.

Gómez Díaz, E., Doherty, P. F., Duneau, D., & McCoy, K. D. (2010). Cryptic vector divergence masks vector-specific patterns of infection: An example from the marine cycle of Lyme borreliosis. *Evolutionary Applications*, *3*, 391–401.

Greiner, M., & Gardner, I. A. (2000). Epidemiologic issues in the validation of veterinary diagnostic tests. *Preventive Veterinary Medicine*, *45*, 3–22.

Hyatt, A. D., Boyle, D. G., Olsen, V., Boyle, D. B., Berger, L., Obendorf, D., … Colling, A. (2007). Diagnostic assays and sampling protocols for the detection of *Batrachochytrium dendrobatidis*. *Diseases of Aquatic Organisms*, *73*, 175–192.

James, T. Y., Toledo, L. F., Rödder, D., da Silva Leite, D., Belasen, A. M., Betancourt-Román, C. M., … Longcore, J. E. (2015). Disentangling host, pathogen, and environmental determinants of a recently emerged wildlife disease: Lessons from the first 15 years of amphibian chytridiomycosis research. *Ecology and Evolution*, *5*, 4079–4097.

Kellner, K. (2015). jagsUI: A wrapper around 'rjags' to streamline 'JAGS' analyses. r package version 1.3.7. http://CRAN.R-project.org/package=jagsUI.

Kery, M. (2002). Inferring the absence of a species – A case study of snakes. *Journal of Wildlife Management*, *66*, 330–338.

Kéry, M., & Schaub, M. (2012). *Bayesian population analysis using WinBUGS: A hierarchical perspective*. Oxford, UK: Elsevier.

Knowles, S. C. L., Wood, M. J., Alves, R., Wilkin, T. A., Bensch, S., & Sheldon, B. C. (2011). Molecular epidemiology of malaria prevalence and parasitaemia in a wild bird population. *Molecular Ecology*, *20*, 1062–1076.

Kriger, K. M., & Hero, J. M. (2006). Large-scale seasonal variation in the prevalence and severity of chytridiomycosis. *Journal of Zoology*, *271*, 352–359.

Kriger, K. M., Hero, J., & Ashton, K. J. (2006). Cost efficiency in the detection of chytridiomycosis using PCR assay. *Diseases of Aquatic Organisms*, *71*, 149–154.

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.

Lachish, S., Jones, M., & McCallum, H. (2007). The impact of disease on the survival and population growth rate of the Tasmanian devil. *Journal of Animal Ecology*, *76*, 926–936.

Laking, A. E., Ngo, H. N., Pasmans, F., Martel, A., & Nguyen, T. T. (2017). *Batrachochytrium salamandrivorans* is the predominant chytrid fungus in Vietnamese salamanders. *Scientific Reports*, *7*, 44443.

Langwig, K. E., Frick, W. F., Reynolds, R., Parise, K. L., Drees, K. P., Hoyt, J. R., … Kilpatrick, A. M. (2015). Host and pathogen ecology drive the seasonal dynamics of a fungal disease, white-nose syndrome. *Proceedings of the Royal Society B*, *282*, 20142335.

Langwig, K. E., Voyles, J., Wilber, M. Q., Frick, W. F., Murray, K. A., Bolker, B. M., … Kilpatrick, A. M. (2015). Context-dependent conservation responses to emerging wildlife diseases. *Frontiers in Ecology and the Environment*, *13*, 195–202.

Lips, K. R., Brem, F., Brenes, R., Reeve, J. D., Alford, R. A., Voyles, J., … Collins, J. P. (2006). Emerging infectious disease and the loss of biodiversity in a Neotropical amphibian community. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 3165–3170.

Lips, K. R., Reeve, J. D., & Witters, L. R. (2003). Ecological traits predicting amphibian population declines in Central America. *Conservation Biology*, *17*, 1078–1088.

Longcore, J. E., Pessier, A. P., & Nichols, D. K. (1999). *Batrachochytrium dendrobatidis* gen et sp nov, a chytridpathogenic to amphibians. *Mycologia*, *91*, 219–227.

Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012). *The BUGS book: A practical introduction to Bayesian analysis*. Boca Raton, FL: CRC Press.

Martel, A., Blooi, M., Adriaensen, C., Van Rooij, P., Beukema, W., Fisher, M. C., … Pasmans, F. (2014). Recent introduction of a chytrid fungus endangers Western Palearctic salamanders. *Science*, *346*, 630–631.

Martel, A., Spitzen-van der Sluijs, A., Blooi, M., Bert, W., Ducatelle, R., Fisher, M. C., … Pasmans, F. (2013). *Batrachochytrium salamandrivorans* sp. nov. causes lethal chytridiomycosis in amphibians. *Proceedings of the National Academy of Sciences*, *110*, 15325–15329.

Miller, D. A. W., Talley, B. L., Lips, K. R., & Campbell Grant, E. H. (2012). Estimating patterns and drivers of infection prevalence and intensity when detection is imperfect and sampling error occurs. *Methods in Ecology and Evolution*, *3*, 850–859.

Muths, E., Corn, P. S., Pessier, A. P., & Green, D. E. (2003). Evidence for disease related amphibian decline in Colorado. *Biological Conservation*, *110*, 357–365.

Olson, D. H., Aanensen, D. M., Ronnenberg, K. L., Powell, C. I., Walker, S. F., Bielby, J., … Fisher, M. C. (2013). Mapping the global emergence of *Batrachochytrium dendrobatidis*, the amphibian chytrid fungus. *PLoS ONE*, *8*, e56802.

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Ritacco, V., Lopez, B., Dekantor, I. N., Barrera, L., Errico, F., & Nader, A. (1991). Reciprocal cellular and humoral immune-responses in bovine tuberculosis. *Research in Veterinary Science*, *50*, 365–367.

Royle, J. A., & Nichols, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology*, *84*, 777–790.

Ruiz-Gutiérrez, V., Zipkin, E. F., & Dhondt, A. (2010). Occupancy dynamics in a tropical bird community: Unexpectedly high forest use by birds classified as non-forest species. *Journal of Applied Ecology*, *47*, 621–630.

Tetzlaff, S., Allender, M., Ravesi, M., & Smith, J. (2015). First report of snake fungal disease from Michigan, USA involving Massasaugas, *Sistrurus catenatus* (Rafinesque 1818). *Herpetology Notes*, *8*, 31–33.

Thompson, K. G. (2007). Use of site occupancy models to estimate prevalence of *Myxobolus cerebralis* infection in trout. *Journal of Aquatic Animal Health*, *19*, 8–13.

Toft, N., Jørgensen, E., & Højsgaard, S. (2005). Diagnosing diagnostic tests: Evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. *Preventive Veterinary Medicine*, *68*, 19–33.

Valkiunas, G., Iezhova, T. A., Krizanauskiene, A., Palinauskas, V., Sehgal, R. N. M., & Bensch, S. (2008). A comparative analysis of micropscopy and PCR-based detection methods for blood parasites. *Journal of Parasitology*, *94*, 1395–1401.

Vredenburg, V. T., Knapp, R. A., Tunstall, T. S., & Briggs, C. J. (2010). Dynamics of an emerging disease drive large-scale amphibian population extinctions. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 9689–9694.

Wake, D. B., & Vredenburg, V. T. (2008). Are we in the midst of the sixth mass extinction? A view from the world of amphibians. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 11466–11473.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: DiRenzo GV, Campbell Grant EH, Longo AV, Che-Castaldo C, Zamudio KR, Lips KR. Imperfect pathogen detection from non-invasive skin swabs biases disease inference. *Methods Ecol Evol*. 2017;00:1–10. https://doi.org/10.1111/2041-210X.12868