

Validation of laboratory tests for infectious diseases in wild mammals: review and recommendations

Journal of Veterinary Diagnostic Investigation
2020, Vol. 32(6) 776–792
© 2020 The Author(s)
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1040638720920346
jvdi.sagepub.com

Beibei Jia,¹  Axel Colling, David E. Stallknecht, David Blehert, John Bingham, Beate Crossley,  Debbie Eagles, Ian A. Gardner

Abstract. Evaluation of the diagnostic sensitivity (DSe) and specificity (DSp) of tests for infectious diseases in wild animals is challenging, and some of the limitations may affect compliance with the OIE-recommended test validation pathway. We conducted a methodologic review of test validation studies for OIE-listed diseases in wild mammals published between 2008 and 2017 and focused on study design, statistical analysis, and reporting of results. Most published papers addressed *Mycobacterium bovis* infection in one or more wildlife species. Our review revealed limitations or missing information about sampled animals, identification criteria for positive and negative samples (case definition), representativeness of source and target populations, and species in the study, as well as information identifying animals sampled for calculations of DSe and DSp as naturally infected captive, free-ranging, or experimentally challenged animals. The deficiencies may have reflected omissions in reporting rather than design flaws, although lack of random sampling might have induced bias in estimates of DSe and DSp. We used case studies of validation of tests for hemorrhagic diseases in deer and white-nose syndrome in hibernating bats to demonstrate approaches for validation when new pathogen serotypes or genotypes are detected and diagnostic algorithms are changed, and how purposes of tests evolve together with the evolution of the pathogen after identification. We describe potential benefits of experimental challenge studies for obtaining DSe and DSp estimates, methods to maintain sample integrity, and Bayesian latent class models for statistical analysis. We make recommendations for improvements in future studies of detection test accuracy in wild mammals.

Key words: detection tests; infectious diseases; sensitivity; specificity; validation; wild mammals.

Introduction

Improved understanding of the ecology of wildlife disease and the role of wildlife in the evolution and transmission of emerging infectious diseases is a core component of modern One Health initiatives. Much of this information is acquired through ecologic and epidemiologic studies in populations, and such studies require accurate and reproducible detection tests to provide information about disease or infection status of individuals and populations. Information gained in these studies is vital for guiding and monitoring health management policies including control of important animal and zoonotic diseases. The concept of a test being fit for a specified purpose (e.g., screening or confirmatory diagnosis) has been endorsed and promoted by the World Organisation for Animal Health (OIE) and others who work on test validation science.^{36,40}

To enhance the harmonization of tests used in international trade, the OIE developed a standardized validation pathway¹²¹ with 4 phases, representing analytical (stage 1), diagnostic (stage 2), among-laboratory reproducibility studies (stage 3), and deployment of tests to multiple laboratories (stage 4). Adherence to the OIE test validation pathway, development of reporting standards for test accuracy studies,^{34,35} and design

standards in some animal species or taxa^{57,77} have improved the utility of studies by better defining source populations, use of structured experimental designs including sampling methods, and accounting for strengths and weaknesses of available tests for a specific testing purpose. The focus of these improvements, as outlined in the OIE Manual of Diagnostic Tests and Vaccines for Terrestrial Animals, has been on infectious diseases of livestock and aquatic animals, but a specific chapter (2.2.7)¹²² is directed at validation strategies in wild animals.

Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, Charlottetown, Canada (Jia, Gardner); CSIRO Australian Animal Health Laboratory, Geelong, Victoria, Australia (Colling, Bingham, Eagles); Southeastern Cooperative Wildlife Disease Study, College of Veterinary Medicine, University of Georgia, Athens, GA (Stallknecht); U.S. Geological Survey, National Wildlife Health Center, Madison, WI (Blehert); California Animal Health and Food Safety Laboratory, University of California–Davis, Davis, CA (Crossley).

¹Corresponding author: Beibei Jia, Department of Health Management, Atlantic Veterinary College, University of Prince Edward Island, 550 University Avenue, Charlottetown, Prince Edward Island, Canada C1A4P3. jbeibei@upe.ca

The large numbers of wildlife species, the diversity of pathogens that infect them, and the emergence of new diseases present unique challenges for test validation. Consequently, detection testing algorithms are often inadequately validated. Diseases in wildlife are frequently novel and can include multiple host species. Moreover, wildlife diseases, especially newly emerging ones, also are often studied inadequately. Compared with livestock, much less is known about the accuracy of detection tests when used in wildlife, and the question remains whether there are significant differences in test accuracy (e.g., diagnostic sensitivity [DSe] and specificity [DSp]) when used in different wild mammal species.

Often it is assumed that test protocols that have been validated for related species can be applied equally in wildlife, for example, bovine tuberculosis (*Mycobacterium bovis*). However, frequently there are flaws in the original test validation studies, with incomplete reporting of relevant animal- and population-level information. Access to adequate numbers and quality of specimens, information about infection status of populations, and limited knowledge of host–pathogen biology, pathogenesis, epidemiology, and population dynamics of naturally occurring infections are frequent limitations in wildlife. Also, inadequate funding may limit the scope and size of validation studies, especially when tests are used in multiple wild mammal species.

Our objectives are to: 1) review studies published between 2008 and 2017 on laboratory test validation with purposes of disease management and conservation in wild mammals, focusing on study design and statistical analysis methods; 2) document strengths, weaknesses, and design trade-offs of test accuracy studies in wild mammals including situations in which animal experiments may add value; and 3) make recommendations for improvements in design and statistical analysis, reporting of validation studies, and interpretation of test results. In addition, 2 coauthors (D.E. Stallknecht, D. Blehert) provide insights through case studies that demonstrate how new pathogens or viral serotypes or genotypes can be detected and testing algorithms changed, and how tests evolve for a new disease once the causative agent has been identified.

Methodologic review

Literature search, appraisal, and review process

Our methodologic review was based on checklist items in the Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA).⁷⁰ Specifically, items in the Methods section of the list were used in the identification, screening, inclusion, and summary of relevant literature, including search, study selection, data collection, and synthesis of results.⁷⁰ We first conducted a systematic search to locate existing studies reporting test accuracy for OIE-listed diseases in wild mammals published between 2008 and 2017, using a standard set of search terms

and disease names (Suppl. Fig. 1; Suppl. Data 1, 2). We characterized each study using a 17-item template (Table 1), which was based on modifications of the 27 items recommended in the Statement for Reporting Studies of Diagnostic Accuracy (STARD).¹⁷

We summarized results of full-text review based on whether the study was related to infection with *Mycobacterium bovis* ($n = 15$; Table 2) or pathogens other than *M. bovis* ($n = 30$; Table 3). We hypothesized that the long history of research involving development and validation of tests for *M. bovis* might contribute to substantial differences between these 2 groups. Historically, test accuracy studies for *M. bovis* focused on disease control and veterinary public health, including identification of infected cattle populations and removal of test-positive cattle. Since ~2010, there has been an increased focus on *M. bovis* in wild mammals because of their possible role as reservoir species as well as conservation concerns. In contrast, test validation studies for non-*M. bovis* diseases were mostly for outbreak investigation or cross-sectional studies involving prevalence estimation. Nineteen diseases were included in the non-*M. bovis* studies. Diseases with more than a single manuscript recording were African swine fever ($n = 4$), *Sarcoptes scabiei* ($n = 3$), *Anaplasma* spp. infection ($n = 2$), brucellosis ($n = 2$), chronic wasting disease ($n = 2$), Johne's disease ($n = 2$), rabies ($n = 2$), and *Toxoplasma gondii* infection ($n = 2$; Table 3).

Limitations of published validation studies

The quality of test validation studies in wild mammals continues to improve, but many tests originally developed for humans and domestic animals have often not been validated adequately to include assessment of whether the test protocols or the designated cutoff value are appropriate for wild mammals.¹²⁰ The inconsistent performance of tests applied in phylogenetically related species or geographical locations remains a challenge for the management and control of infectious diseases in wild mammal populations.³⁰ We provide an overview of major limitations in the study design and results and inferences based on our full-text reading of papers, including source populations, sampling scheme, and interpretation of results.

OIE test validation chapters and related studies^{34,36,40,49,121,122} were cited in only 2 of 15 *M. bovis* papers and 10 of 30 non-*M. bovis* papers. These sources of information are best used during the design stage to ensure that the important elements that affect methodologic quality are not overlooked. Adherence to these principles will influence the quality of studies in terms of sampling, statistical methods, and inference.

Study design

Lack of description of the source population. Only 9 of 30 non-*M. bovis* studies provided information about the source population, compared with 9 of 15 *M. bovis* studies. For many test validation studies in wild mammals, the lack of

Table 1. Items used in the full-text appraisal of 45 papers including the relevant item numbers in the Standards for Reporting of Diagnostic Accuracy (STARD)17 Statement. Results are summarized in Tables 2 and 3.

Category	STARD item	Interpretation	STARD no.
Publication identifier	1 Reference and author	Family name of first author	2, 3,4
	2 Year of publication		2, 3,4
Introduction/abstract	3 Purpose	Study objectives	2, 3, 4
Methods			
Study design	4 Disease/pathogen	Common and Latin names	
	5 Species	Common and Latin names	
	6 Source population	Population from which mammals were sampled	
	7 Source population matches purpose	Whether source population matches study purpose	7, 8, 9
Test methods	8 Sample size		18
	9 Specimen	Specimen types (e.g., tissues, feces, serum)	
	10 Index test	Test(s) validated in the study	10a
Analysis	11 Reference standard(s)	Detailed description (not needed if latent class models were used)	10b, 11, 12b
	12 Statistical methods	Methods for estimating or comparing measures of diagnostic accuracy	14
Test results	13 DSe	No. of test positives/no. of true positives	23, 24, 25
	14 No. of true positives	Denominator for DSe	23, 24, 25
	15 DSp	No. of test negatives/no. of true negatives	23, 24, 25
	16 No. of true negatives	Denominator for DSp	23, 24, 25
Discussion	17 Comments	Whether background of target population was provided; how statistical uncertainty and generalizability were addressed	12, 20, 26, 27

DSe = diagnostic sensitivity; DSp = diagnostic specificity.

source population information was not addressed, raising questions about their internal validity. When source population information was missing, it was not easy to assess whether studied animals could be used to make inferences about target populations (in which tests will be used in future) or whether sampled animals matched the target population defined in the study design. For example, calculation of DSe and DSp from clinically affected animals would be of minimal value if the assay was proposed to be used for certification of pathogen freedom in populations with no evidence of clinical diseases.

Lack of representativeness of the source population. Uncertainty about the temporal dynamics of wild mammal populations and the baseline frequency or threshold of infectious diseases in these populations influences the representativeness of the source population compared with the target population. Only a few non-*M. bovis* studies discussed the representativeness of source populations, including the cross-boundary collaboration between Canada and the United States to address chronic wasting disease (CWD) in white-tailed deer¹¹¹ and the validation of tests for *Trichinella* infections in wild boar in Sweden.²¹

Additional examples of lack of representativeness between source and target populations can be also be found in other studies,⁷⁶ such as the use of a mix of experimental and field infection data, and for analysis of pooled data from wild and farmed mammals. In these instances, it is difficult to address how well the sampled animals matched the target

population defined in the research question, which is a metric of internal validity of the study.⁹⁵ For test validation studies using multiple species, estimates of DSe and DSp were sometimes not reported separately for each species.^{28,97}

Sampling scheme: sample selection, quality, and quantity

Sampling methods used in wild mammal populations are usually non-probability methods, such as convenience⁷⁶ or opportunistic sampling⁹⁰ because animals are often darted or trapped for sample acquisition. For example, the potential omission of follow-up samples for validation testing may result in over- or underestimation of DSe because of a lack of willingness of owners to allow an autopsy of *M. bovis* screening test-positive elk in a free-ranging herd.⁷⁶ The estimated DSp might not be representative for disease-free populations if DSp was only based on test-negative animals in infected populations.⁹⁰

Sufficient sample size can be a challenge for providing robust estimates of DSe and DSp.¹⁹ Few studies justified their selected sample size, although most reported 95% confidence intervals (CIs) for DSe and DSp. Selection of an appropriate sample size for estimation of these parameters depends on their expected values, the desired error margin (e.g., $\pm 5\%$), and whether specimens are of known or unknown infection status.¹²¹ Limited numbers of known positive samples are a common limitation that influences the precision of DSe, and this issue is not unique to wild mammals.¹⁸ Logistics is a

Table 2. Summary of selected test validation studies ($n = 15$) for *Mycobacterium bovis* in free-range, captive, and farmed wild mammals published between 2008 and 2017.

Reference no.	Publication year	Species	Specimen type	Index test(s)	Reference standard(s)	Source population available/population matches purpose	Comments on study design and inference*
67	2008	Badgers, white-tailed deer, brushtail possums, wild boar	Serum	MAPIA, rapid test	Culture	Yes/partial	
26	2009	Wild meerkats	Serum, submandibular lymph node	MAPIA, rapid test	Culture	Yes/partial	No details about selection of animals. Repeated samples from same individual animals. Used Bayesian latent class analysis.
78	2009	White-tailed deer	Whole blood, serum	MAPIA, rapid test, immunoblot, ELISA	Culture	Yes/no	No. of positive and negative results not specified.
39	2009	Fallow deer, roe deer, red deer	Serum	Rapid test	Culture	No/no	Infection can be masked by latent or subclinical infection. Low positive predictive value could be a result of the stage of infection and cross-reactions.
45	2010	Wood bison	Serum	FPA, CFT, MAPIA, rapid test	Culture and histopathology	Yes/no	All animals were tested. Challenge of detecting infected individuals with antemortem tests was described.
76	2012	Elk, white-tailed deer, reindeer	Serum	Stat-Pak, MAPIA, DPP VetTB assay	SCT	No/unsure	Focused on DSP and estimate for DSe was from a limited number of samples. DSe determined using serologic test could be lost using skin test. Convenience sampling was used.
8	2012	Fallow deer	Serum	DPP VetTB assay (DPP ₅₋₅), bPPD ELISA	Culture	No/no	Mix of wild and farmed deer for both confirmed infected and non-infected subpopulations.
79	2013	Black bears, bobcats, coyotes, opossum, raccoon, red fox, etc.	Historical database	Gross lesions, histopathology, acid-fast bacteria	Culture	No/no	Secondary data analysis. Pooling of results from multiple species might bias estimates, thereby limiting inference.
68	2013	White-tailed deer	Serum	DPP VetTB assay	Culture	No/no	Mix of experimental infections, wild, and farmed deer.
116	2014	White-tailed deer	Serum	Ethanol vortex ELISA	Culture	No/no	Mix of experimental infections, wild, and farmed deer. OIE validation chapters or related studies cited.
99	2014	Elk	Serum	LST, FPA, Cervid TB Stat-Pak	Culture	Yes/yes	Opportunistic sampling ⁶⁴ . Bias uncorrected. Used classical statistical tests and latent class analysis for test validation.
114	2016	African buffalo	Serum	SICTT, IFN- γ release assays (IRA), ELISA	Culture	Yes/unsure	Relative DSe was reported. Limited sample size but 95% CI reported. OIE validation chapter or related studies cited.
90	2016	Warthogs	Serum	Indirect PPD, ELISAs, TB ELISA-VK, DPP VetTB assay	Culture	Yes/unsure	Opportunistic sampling. Limited sample size, but 95% CI reported.
3	2017	Eurasian wild boar	Oral, nasal, fecal swabs	PCR	Culture	Yes/unsure	Limited sample size with wide CI.
107	2017	European badger	Feces	IMS-IFD, IMS-qPCR, IMS-culture	Culture	Yes/partial	Two sets of samples might represent different source populations.

CFT = caudal-fold test; CI = confidence interval; Culture = mycobacterial culture; DPP = dual-path platform; ELISA = enzyme-linked immunosorbent assay; DSe = diagnostic sensitivity; DSP = diagnostic specificity; FPA = fluorescence polarization assay; IMS-IFD = immunochromatographic lateral flow device; IMS-qPCR = immunomagnetic separation quantitative PCR; IRA = gamma-interferon release assay; LST = lymphocyte stimulation test; MAPIA = multi-antigen print immunoassay; PPD, bPPD = purified protein derivative, bovine purified protein derivative; rapid test = a lateral-flow rapid test; SICCT = single intradermal comparative cervical test; Stat-Pak = lateral flow immunoassay rapid test. For interpretation of convenience and opportunistic sampling, refer to the Results.
 *Whether OIE chapters or related studies^{34,36,40,49,121,122} were cited; this indicates whether the paper has followed principles or methodologic guidelines for test validation in wildlife mammals.

Table 3. Summary of selected test validation studies ($n = 30$) for diseases other than *Mycobacterium bovis* in free-range, captive, and farmed wild mammals published between 2008 and 2017.

Reference no.	Year	Disease or pathogen	Species	Specimen type	Index tests	References test(s)	Source population/ matches purposes	Principle cited*/ statistical methods
58	2008	Classical swine fever virus	Wild boar	Organs (tonsil, lymph nodes, ileocecal valves, kidney, spleen), blood, viral culture	2 RT-PCRs	Virus isolation	Mix of PT panels and field/no	Yes/none
97	2008	<i>Anaplasma</i> spp.	Wild ruminants	Serum	ELISA	PCR, IFA	Retrospective samples of mixed species of wild ruminants/no	No/ROC
125	2009	<i>Trypanosoma cruzi</i>	Raccoon, Virginia opossum	Serum	Chigas Stat-Pak	IFA	Archived samples, samples from experimental infections/no	No/point estimate with CI
29	2010	<i>Anaplasma phagocytophilum</i> , parvovirus	Fishers, gray foxes	Fecal samples, whole blood	SNAP-PARVO RDT	PCR, IFA	None/NA	No/point estimate with CI
96	2010	<i>Brucella abortus</i>	Elk	Serum	Card test, Rivanol, RAP, western blot	Bacterial culture	Some provided by collaborators with known infection/no	No/point estimate with CI
112	2011	African swine fever virus	Wild boar, domestic pigs, bush pigs, and ticks	Whole blood, serum, tissues, ticks	Real-time PCR with internal control	Virus isolation	None/NA	No/point estimate
6	2012	Chronic wasting disease: PrP ^{Sc} ^{WD}	Rocky Mountain elk, white-tailed deer, mule deer, roe deer, red deer, fallow deer	Brainstem (obex region) and lymph node tissues	4 ELISAs, western blots, lateral flow test	IHC	None/NA	No/only analytical Se and Sp reported
111	2012	Chronic wasting disease	White-tailed deer	Brainstem, lymph node tissues	IHC	Rectal biopsy test results	3 Canadian and 1 U.S. herd/ confined to case herds with complete information	No/random-effects logistic regression model; Forest plots for DSe
126	2012	Rabies virus	Raccoon dogs, cattle, dogs	Brain	Virus isolation, RT-PCR, RIDA	Fluorescent antibody test	None/NA	No/point estimate
80	2013	Henipavirus	Fruit bat	Plasma	Multiplexed microsphere binding assay	Virus neutralization	Opportunistic sampling/mixed samples	Yes/ROC, Bayesian
84	2013	MAP infection	Elk, bison, caribou	Serum	ELISA	Bacterial culture	Animals of known infectious status/mixed samples	Yes/point estimate, ROC
106	2013	Infestation of <i>Notodres catt</i> (not OIE-listed)	Bobcats	Fecal samples	PCR	Clinical sign, skin scraping, fecal flotation	Bobcats in the preserve area where no significant population fluctuation/opportunistic sampling	No/point estimate
118	2013	Rabies virus	Red fox, racoon dog	Serum	BioPro ELISA	FAVN	Field samples and experimental infection/no	No/specificity concordance reported
1	2014	<i>Toxoplasma gondii</i>	Gray seal	Serum	ELISAs, western blot	Not applied	None/NA	Yes/relative values reported
21	2014	<i>Trichinella</i> infection	Wild boar	Serum	ELISAs, western blot	Artificial digestion	Low- and high-risk groups/no	No/only kappa
83	2014	MAP	Fallow deer	Serum	ELISAs	Histopathology, culture	Samples from a 500-deer population/partial	Yes/ROC, kappa

(continued)

Table 3. (continued)

Reference no.	Year	Disease or pathogen	Species	Specimen type	Index tests	References test(s)	Source population/ matches purposes	Principle cited*/ statistical methods
13	2015	<i>Sarcoptes scabiei</i>	Experimentally infected, farmed, and wild rabbits	Serum	ELISA	Skin scraping	Wild rabbits from Mallorca Island/no	No/point estimate
31	2015	African swine fever virus	Wild boars	Tissues, bone marrow, whole blood, serum	PCRs, ELISA	Virus isolation	Infection-status known animals/no	No/point estimate
38	2015	Brucellosis	African buffalo	Serum	ELISA, RBT, CFT	RBT, CFT	Herd from Kruger National Park/partial	Yes/Bayesian modeling
41	2015	<i>Sarcoptes scabiei</i>	Wild boars	Serum	ELISA	Skin scraping	Wild boar from Petite Pierre National Reserve/repeatable sampling on same animals	Yes/point estimate
117	2015	<i>Toxoplasma gondii</i>	Wild boars	Serum	ELISA	None	Wild boar population in Sweden/random selection from national surveillance program	Yes/Bayesian latent class analysis
23	2016	<i>Leishmania infantum</i>	Rabbit, hares	Serum, skin	IFATs, PCR	None	Lagomorphs in 2 areas of the region of Madrid/partial	Yes/Bayesian approach
48	2016	Tick-borne encephalitis virus	Wild rodents	Serum	ELISA, PCR	Neutralization test	Mixed samples/no	No/point estimate
86	2016	<i>Sarcoptes scabiei</i>	Iberian ibex	Serum	ELISA	Categorization of scabietic skin lesion	Iberian ibexes from Sierra Nevada Natural and National Parks/convenience sampling	No/ROC curve
94	2016	African swine fever virus	Wild boar	Serum	Lateral flow assay	PCR, antigen-detection ELISA	Mixed samples/no	No/point estimate
44	2016	Bluetongue	Bovine/ovine/caprid	Serum	Lateral flow assay, PCR	Antibody-ELISA	Mixed samples/no	Yes/point estimate
12	2017	African swine fever virus	Wild boar	Serum, lung	Pen-side rapid test	IPMA, ELISA	Wild boar on island of Sardinia, Italy/yes	No/ROC curve
28	2017	Epizootic hemorrhagic disease	Impala, dromedary camel	Serum	ELISA	Virus isolation	Infectious-status known samples/no	No/point estimate
59	2017	Bovine spongiform encephalopathies	Macaques, human, cattle	Brain	RT-QuIC	Western blot	None/NA	No/point estimate
110	2017	Hemorrhagic septicemia	Asian elephants	Serum	ELISA, IHA	None	Archived samples/no	Yes/Bayesian approach

CFT = complement fixation test; FAVN = fluorescent antibody virus neutralization test; IHA = indirect hemagglutination assay; IHC = immunohistochemistry; IPMA = immunoperoxidase monolayer assay; NA = not applicable (when source population is not stated, it is not appropriate to assess whether the source population matches the study purpose); MAP = *Mycobacterium avium* subsp. *paratuberculosis*; PT = proficiency test; RBT = Rose Bengal test; RIDA = rapid immunodiagnostic assay; RT-QuIC = real-time quaking-induced conversion assay.

*Whether OIE chapters or related studies^{34,35,46,49,71,73,74} were cited; this indicates whether the paper has followed principles or methodologic guidelines for test validation in wild mammals.

widely recognized factor for limited longitudinal follow-up of test positives, especially when multiple populations are studied. Repeated measurements to demonstrate temporal variation in serum antibody responses might not be realistic given the difficulty of resampling the same animals.

Sample quality often negatively impacts test accuracy for both direct detection assays (e.g., culture and PCR) and indirect assays (e.g., serologic and gamma-interferon tests). Storage, transportation, and handling of samples can also adversely affect test performance. Processing (freeze–thaw) and transportation were reported with possible decreased viability of the organism in bronchoalveolar lavage samples for culture of *M. bovis* from African lions (*Panthera leo*), especially those specimens collected from latently infected animals with a low initial count of bacilli. Low DSe for diagnosing individual wild mammals shedding *M. bovis* can happen with the combined conditions of small volume sample and temperature fluctuation during storage and transport.²⁶ Use of FTA cards to ensure long-term sample quality is described below in our study. None of the 45 papers provided data about effects of sample quality on DSe and DSp estimates or addressed methods to improve sample quality. Acceptance criteria for sample suitability facilitate a minimum level of quality and are a requirement in accredited laboratories (see also recommendations at the end of our study).

Results and inferences

Insufficient discussion of bias induced by heterogeneity of the source population. Given the long latent period of chronic pathogens, such as *M. bovis*, it is logistically difficult to track relevant information related to the population structure and temporal dynamics of infection, which is best characterized if animals can be individually identified, which is rarely done.³⁹ Furthermore, understanding basic population structure and dynamics is not possible in most cases, and if possible, would require a long-term commitment of substantial resources to allow repeated testing of surviving animals.

Insufficient consideration of impact of infection stage, disease progression, and cross-reacting infectious agents on test results. Data interpretation and reporting of DSe and DSp need to consider that spatial-temporal dynamics of susceptible and infected wildlife vary among geographic areas.¹⁰⁵ Furthermore, in addition to confounding factors such as age, sex, pregnancy, vaccination, stress, and nutritional status, test accuracy can be affected by the interaction between the pathogen of interest and the diversity of cross-reacting infectious agents in the host species.¹⁴ None of the 15 *M. bovis* papers considered comorbidity issues that may lead to misclassification bias because of cross-reactions from coinfecting *Mycobacterium* spp. or reduced immunologic response associated with viral infections such as feline immunodeficiency virus (FIV). FIV coinfections negatively affect *M. bovis* test

performance in African lions^{54,108} because of effects on both antibody and cell-mediated immune responses.^{82,88} Test validation results should be interpreted with respect to how missing information, including data on coinfections, may influence the proportion of positive and negative results.^{8,99}

General progress in test validation for wild mammal diseases

Based on the full-text reading of the 45 manuscripts in our literature search, we describe how researchers have made progress embracing challenges in the validation of fit-for-purpose assays for wild mammal species, especially in the areas of study design and statistical analysis (Table 1).

Study design

Samples from reference animals of known infection status. Given the logistical challenges of collecting samples from wild mammals, strategies are needed to acquire sufficient numbers of positive and negative reference samples for estimation of DSe and DSp. This is a challenge for laboratory test evaluation in most wildlife diseases, especially for non-*M. bovis* studies. Utilization of samples of known infection (archived) status was done in 20% of studies for diseases including African swine fever, *Brucella abortus*, epizootic hemorrhagic disease, hemorrhagic septicemia, Johne's disease, and *Trypanosoma cruzi*. However, only 3 *M. bovis* studies^{67,76,107} and 3 non-*M. bovis* studies^{84,96,110} used field samples for validation of novel or existing tests in addition to the use of samples from reference animals or experimentally infected wild mammals. The latter sample types often result in overestimation of DSe and DSp.

Statistical analysis

Most selected papers incorporated estimates of uncertainty in test performance. For example, studies of *M. bovis* reported CIs for both DSe and DSp, with a few exceptions because of the sample sizes available.^{8,76} Postmortem methods were shown to have DSp close to 100% with narrower CIs but low DSe with wide CIs (Suppl. Figs. 1, 2; Suppl. Data 1, 2). Some studies provided predictive values for a range of expected prevalence of infection in the target population.^{39,45,96,97,99,106}

Bayesian latent class model (LCM) has been used increasingly for estimation of DSe and DSp for wild mammal diseases to account for imperfect reference tests. The use of LCM has been acknowledged by the OIE as a valid method for analysis of test accuracy data (stage 2 of the OIE pathway) in the context of assessing a test's fitness for a defined purpose (e.g., surveillance, confirmation of infection or disease, and prevalence studies).^{121,122} Two *M. bovis* studies used LCM for the estimation of DSe and DSp with 95% probability intervals.^{26,99} Four non-*M. bovis* studies reported LCM as the statistical analysis method.^{23,38,80,117} A cost-effec-

tive ELISA for screening for brucellosis in African buffalo (*Syncerus caffer*) was developed with determination of cut-off value based on the use of a Bayesian LCM.³⁸ Similar methods were also applied in henipavirus assays in African bats.⁸⁰

The quality of papers using Bayesian LCM for the period from 2008 to 2017 is variable. For *M. bovis*, a 3-test in 1-population model was used for comparison of DSe and DSp of the caudal-fold test (CFT) and 2 serologic tests, a lateral flow rapid test (RT) and fluorescent polarization test (FPA), in 212 free-ranging wild wood bison (*Bison bison athabasca*), with the latter tests assumed to be conditionally dependent.¹⁵ Given that this model is non-identifiable (i.e., there is no unique set of values for DSe and DSp given the joint test results), prior information on at least 2 parameters was necessary. The authors used prior information on 5 parameters, which should have ensured model identifiability, if at least 1 test was conditionally independent of the other 2 tests. Model results for the FPA indicated that the assay was less accurate than tossing a coin (i.e., worthless). OpenBUGS code was not provided in the paper, which would have allowed readers to identify possible reasons why such results were obtained. In another study, the accuracy of 2 new serologic tests was compared, multi-antigen print immunoassay (MAPIA) and a lateral-flow rapid test (RT), with results of culture of tracheal washes in 110 meerkats.²⁶ Three approaches for the statistical analysis, including a Bayesian LCM, were used. The joint counts for all test result combinations of positive and negative results were presented in Table 2 of their paper, and the Bayesian estimates were realistic, with culture being the most specific but least sensitive test. Code and convergence testing were not reported. In a follow-up study, the accuracy of mycobacterial culture, gamma-interferon assay, and a commercial lateral-flow immunoassay (Stat-Pak; DPP VetTB assay for cervids, DPP VetTB assay for elephants; Chembio Diagnostic Systems, Medford, NY) were compared in 305 badgers (*Meles meles*).²⁷ A Bayesian model of 3 conditionally independent tests in 1 population was used. Adequate sensitivity analysis was done, and sources of prior information (expert opinion and published studies) were well justified. However, a code was not provided, therefore it is unclear whether the values for use of tests in parallel were done within the Bayesian code and whether uncertainty in the estimates was captured. In conclusion, all 3 papers would have benefited from adherence to current STARD-BLCM (Bayesian LCM) standards.⁵⁶

Reporting and discussion

Use of multiple tests for interpretation of results. Combinations of tests that are run simultaneously or sequentially and interpreted in series or in parallel are used for detection of chronic bacterial infections caused by *M. bovis*, *M. paratuberculosis*, or *Brucella abortus*. Measuring different aspects of immunity (cellular- vs. antibody-based) for *M. bovis* will

likely result in a higher combined DSe for detecting infection than 2 tests that measure the same immune response.⁶⁸ Even with antemortem tests, several researchers noted limitations of the validation results and challenges in detecting *M. bovis*-infected animals. Seroreactive animals were not culture-positive for *M. bovis*.¹¹⁴ Animals testing negative by rapid tests (e.g., Stat-Pak) were only confirmed by autopsy if they underwent slaughter inspection.⁸ Inconsistencies among serologic tests existed, and those negatives were shown to be culture or histopathology positive for tuberculosis, which indicated that those antemortem tests cannot definitively detect *M. bovis* infection but in return are likely to be highly specific.⁴⁵ In some studies, postmortem descriptions of sampled animals were also provided by incorporating other information (e.g., descriptions of gross lesions and histopathology) and antemortem test results (e.g., mycobacterial culture) from other animals in the source population.⁶⁸

Combinations of tests have been used for early detection of highly contagious diseases, such as African swine fever (ASF). Compared with other serologic tests, the immunoperoxidase test (IPT) has higher DSe for antibody detection earlier in the serologic response and can thus be used to test tissues of dead wild boars to determine whether the animals were exposed and/or infected with ASF virus (ASFV). The combination of an antibody test (IPT) and antigen-detection test (Universal Probe Library, UPL-PCR) was reported as the most reliable method for detecting ASFV during epidemics in Europe in 2014.³¹ In another study, pen-side antigen and antibody tests used in combination were proposed for cost-effective monitoring of ASF.¹² ASFV-specific nucleic acid was found reliably using enrichment procedures in seized pork products in Australian airports (pers. comm., Jianning Wang, 9 Nov 2019, <https://www.weeklytimesnow.com.au/agribusiness/african-swine-fever-found-in-pork-products-at-australian-airports-and-in-mail/news-story/d78eb71cd-d6119a5e97d28d069549e39>).

Another example is the detection of *Mycobacterium avium* subsp. *paratuberculosis*, the causative agent of paratuberculosis (Johne's disease). Serum ELISAs for paratuberculosis detection in fallow deer (*Dama dama*) were evaluated, and their predictive values were interpreted together with the presence or absence of histologic lesions.⁸³ In the same study, ELISA-based results were compared pairwise to results from other testing techniques, including Ziehl-Neelsen staining, quantitative PCR, and bacterial culture. These multi-test comparisons for paratuberculosis indicated that pathogen exposure likely triggered antibody production without progressive infection.⁸³ Thus, ELISA can offer a low-cost screening tool in paratuberculosis surveillance to segregate fallow deer populations into infected and non-infected groups.⁸³

Improved reporting of pooled data from multiple species or multiple target populations. In an attempt to match the target population, results of testing from subpopulations were

reported together with those of pooled data (i.e., mixed species samples from a source population with different infection histories).^{8,79} In addition to the pooled estimate, separate estimates for each species were provided to mitigate concern about differences in susceptibility, spectrum of disease, age and sex ratio, or phylogenetic distance from the domestic species for which the index test was developed.⁸⁰ For example, rapid tests developed for detecting *Anaplasma phagocytophilum* and parvovirus were validated and reported for fishers (*Martes pennanti*) and gray foxes (*Urocyon cinereoargenteus*), respectively.¹²⁵ Additional examples can be found for *M. bovis* papers, including serologic assays for detection of *M. bovis* in wild ruminants,⁷⁶ and ELISA for detection of the organism in cattle and deer.¹¹⁶ Another study addressed the use of multiple species to establish cutoff values for free-ranging bison and caribou when no reference samples were available and infection status of the tested animals was not known.⁸⁴

Case studies: test validation of endemic diseases in wild mammals

Before proposing general strategies to address the aforementioned challenges, we describe specific problems of test validation for 2 endemic diseases in North America for purposes of disease management. These examples demonstrate that validation is an on-going process as pathogens change, testing technology evolves, and understanding of host–pathogen–environmental interactions in wild mammals improves. The second case of white-nose syndrome (WNS) in hibernating bats may set a precedent and establish an approach for broader harmonization and professionalization of laboratory testing for wildlife diseases.

Hemorrhagic disease in deer

Hemorrhagic disease (HD) is an important mortality factor affecting white-tailed deer (WTD) in North America and is caused by both bluetongue virus (BTV) and epizootic hemorrhagic disease virus (EHDV); mortality associated with BTV and EHDV infections has also been documented in other wild ungulate species, including mule deer, pronghorn antelope, and elk.⁴⁶ The interpretation of tests as applied to field-acquired HD is complicated by multiple and genetically related viruses (BTV and EHDV), multiple serotypes of viruses (at least 6 BTV and 3 EHDV serotypes in North America), diversity of ungulate species that are susceptible to this disease, and extreme variability in clinical outcome. Tests to detect BTV and EHDV include virus isolation in embryonated chicken eggs and cell culture, and various reverse-transcription PCR (RT-PCR) and RT-qPCR protocols.^{46,55,123} In addition, serologic techniques, including virus neutralization and ELISA, are available to detect prior or current exposure to the viruses and are most often used in surveillance or other epidemiologic studies. When these assays are assessed under

controlled experimental conditions in young naïve WTD, it is not uncommon to achieve 100% estimates for both DSe and DS_p depending on stage of infection.⁸⁵ However, currently, there are no reliable estimates of either DSe or DS_p when applied to North American ungulates naturally infected with either BTV or EHDV.

Test performance of virus detection techniques may vary over the course of infection. Both BTV and EHDV can be associated with a long duration (up to 50 d) of viremia, but duration is highly variable between strains and possibly serotypes.^{46,85} Although estimates of BTV or EHDV RNA persistence are not available for wild ungulates, it is likely that this also is prolonged as found with BTV in cattle (up to 160 d) and sheep (up to 89 d).⁵³ Long-lasting naturally occurring reductions in antibody titer occur over time, and cross-reactive antibodies between EHDV and BTV serotypes may cause false-positive results.⁵³ The extent of this cross-reactivity may broaden with time as a result of exposure to multiple BTV and EHDV serotypes and possibly to a wider range of microorganisms and other nonspecific substances.⁵³ Hence, cross-reactivity may vary greatly between age cohorts within the study population, and geographically, between populations in BTV- and EHDV-endemic and epidemic areas.^{91,103,104}

With cases submitted from the field, sample quality also is an important consideration given that most submissions are initiated by the public and occur in late summer and early fall when environmental temperatures are often high. To date, there are no estimates available related to potential loss of infectivity or loss of detectable RNA associated with sample quality. Similar issues also arise with sera collected from hunter-killed deer because blood collected from body cavities may be of questionable quality.

With HD, as with many wildlife diseases, DSe or DS_p estimates are complicated by the complexity associated with multiple host species, multiple viruses, and highly variable clinical and epidemiologic patterns over the range of these hosts and viruses. The duration and levels of infectious virus, viral RNA, and antibodies will affect DSe and DS_p, but unfortunately, current experimental data on the extent of virus and host variability is inadequate to accurately estimate potential effects on test accuracy.

White-nose syndrome in hibernating bats

WNS caused the most precipitous decline of a North American wild mammal population in recorded history.⁴ Spread of the disease in the northeastern United States is described elsewhere.⁷ The white-and-gray powdery substance observed on muzzles, wings, and tail membranes of affected bats had the gross appearance of a fungus, which was subsequently confirmed by direct microscopy and histopathology.⁷ Culture conditions (initially 4°C) that replicated skin temperatures of hibernating bats were needed for successful isolation of the causative agent, a novel psychrophilic (cold-loving) fungus that was ultimately named *Pseudogymnoascus*

destructans.^{37,72} The fungus has been shown to be the primary causative agent of WNS through fulfillment of Koch postulates.⁶⁴

The initial laboratory-based assessment of skin samples from 117 bats showed a strong association between the presence of *P. destructans* and histologic skin lesions.⁷ Subsequently, to facilitate consistent analyses across multiple diagnostic laboratories, histologic criteria for WNS were established.⁷ Also, to enable more rapid detection of the fungus, a conventional PCR-based method was developed to detect *P. destructans* based upon amplification of a 624-nt fragment of the internal transcribed spacer (ITS) region of the ribosomal RNA (rRNA) gene.⁶³ This test was validated using skin samples from bat carcasses that were previously identified as either positive ($n = 48$) or negative ($n = 32$) for WNS by histology and had 100% DSp and 96% DSe for detection of *P. destructans*. However, when this method was subsequently used to analyze environmental samples, such as soil or sediment collected from bat hibernation sites, the test had lower DSp than when previously validated using a subset of skin samples from infected bats that harbored a less complex background of fungal species and that was effectively enriched for *P. destructans*.

Sequence-based analyses of the PCR amplicons from environmental samples revealed a previously unknown diversity of closely related fungal species that cross-reacted with conventional PCR primers,⁶³ but those were collectively still distinguishable from *P. destructans* by single nucleotide polymorphisms following DNA sequence analysis. However, as genetic differences between *P. destructans* and closely related fungal species are often subtle, sequence-based comparisons must be conducted precisely and require a higher-level understanding of the nuances of GenBank and BLAST. Thus, efforts were undertaken to develop an improved PCR test, based on TaqMan procedures, which would more robustly distinguish *P. destructans* from near-neighbor fungal species without the need for follow-up sequence-based analyses of PCR amplicons.

To address this problem, it was necessary to characterize a less conserved molecular marker, the rRNA gene intergenic spacer (IGS) region, from *P. destructans* and closely related fungal species, to facilitate consistent identification of *P. destructans* without yielding false-positive results.⁶⁵ Following analysis of the resulting IGS region sequence data from 172 fungal isolates, a real-time TaqMan PCR test was developed and validated based on analysis of skin samples from bat carcasses that were previously identified as either positive ($n = 49$) or negative ($n = 42$) for WNS by histology.⁷³ This assay was shown to have perfect DSe and DSp for *P. destructans*, additionally exhibited no cross-reactivity when tested for amplification of nucleic acid extracts from 54 fungal isolates designated as near neighbors of *P. destructans*, and was considered the most accurate PCR-based assay for detecting the organism. With the availability of robust techniques to culture *P. destructans*,⁶³

to identify the presence of fungal DNA by real-time PCR⁷³ and to identify lesions diagnostic for WNS,⁷¹ the final step was to develop a case definition for WNS based on test results combined with field observations to define infection by *P. destructans* and WNS (see case definitions for WNS).¹¹³

The approach for developing test procedures in response to the emergence of WNS has been a multi-step process completed with necessary urgency. The steps have included identification and cultivation of the causative fungus, establishment of histologic criteria for WNS, development of a robust real-time PCR test, and creation of a case definition for infection by *P. destructans* and for WNS. Despite these advances, challenges remain. For example, there are now several PCR tests for *P. destructans* described in the peer-reviewed literature that differ in DSe and DSp,^{16,63,73,98} and which may be applied without discrimination by independent laboratories unfamiliar with strengths and weaknesses of each published test. This presents challenges for standardization or harmonization of testing outcomes from independent laboratories and increases the chances for reporting false-negative or false-positive results. Efforts are now underway to promote harmonization of molecular testing among a voluntary network of laboratories performing PCR-based assays for *P. destructans*.

Strategies at different stages of test validation to improve study design, sample integrity, test methods, and statistical approaches

In the following sections, we elaborate on how test validation has been improved for infectious diseases of wild mammals. We address the use of experimentally- versus field-infected animals, sampling methods for improving long-term sample quality, and latent-class statistical methods for analysis of test-accuracy data from populations of unknown infection or disease status.

Study design: trade-off between experimental infection studies and samples from naturally occurring infections

One of the challenges described in the introduction of our study is the use of samples from naturally occurring infections. Experimental infection studies may address this issue when infection is induced in the wildlife species of interest or in a suitable model species (e.g., ferrets), which is then monitored to acquire relevant clinical data and samples. Strengths of these studies include: 1) suitable controls that can be incorporated in the design; 2) frequent, even continuous, monitoring to support improved clinical information for collection of samples; and 3) the potential to determine incubation periods.

Most experimental infection studies are designed to understand pathogenesis and can be useful to determine the host–pathogen biology of a hypothetical host species. However, appropriate experimental designs also allow estimation of DSe in anticipation of, and in preparation for, pathogen incursion events.⁵ The added value of including estimation of assay performance in infection studies allows for the ability to adequately sample animals to provide quality samples for evaluation of the accuracy of multiple tests, the opportunity to obtain standardized positive-control samples for future field studies, and the ability to evaluate DSe and DSp within a defined timeline.

Host–pathogen biological factors that may impact test accuracy, and that can be determined from infection studies, include the organ and tissue types that support infection (pre-dilection sites), routes of pathogen shedding, kinetics and timing of the pathogen replication cycle, nature and magnitude of immune responses, and presence of related pathogens within the population. These data can also be determined from the study of natural field cases, although such studies may not have the same precision because of the lack of case numbers, quality of samples, and quality of related case information. Nevertheless, such field case information is critical to interpreting results from host–pathogen studies because it helps to verify that experimental infection studies are representative of natural infections. For further details about sample size considerations, readers are referred to epidemiology texts.²⁵

There are different ways in which data on host–pathogen biology can be acquired within animal infection studies. These include clinical monitoring, autopsy and histologic examination, and other pathology assays. Such data help to define each infection case, allowing characterization of norms and variations of the infected captive population, and comparison with natural cases. Test accuracy metrics, such as DSe and DSp, should be interpreted in light of these findings. For example, pteropid bats, which are known to be maintenance hosts for henipaviruses, usually shed only limited virus and develop low levels of antibody following experimental infection with henipaviruses,⁴³ and this may be the result of a limitation of sampling and monitoring protocols in these poorly understood host–pathogen systems.¹⁰⁰

However, experimental infection studies may also have limitations. The acquisition of wildlife is usually complex, resource-intensive, and sometimes unsafe for the animals and their handlers. The most favorable option is to raise animals in captivity within breeding colonies, therefore providing opportunities for acclimatization and training to the conditions and needs of the study. However, setting up these colonies is generally not practical, is always resource-intensive, and rarely can it be justified solely for purposes of studying a disease of interest. Additionally, captive-bred animals will not always be representative of the natural populations and the range of cross-reacting agents to which free-ranging animals may have been exposed.

On the other hand, using captive animals from the wild may be a short-term option, but needs to account for the stress and safety of capturing animals and in removing them from their natural environmental and social settings. Stresses related to capture and captivity may alter pathophysiological responses^{11,24,89} and may, in turn, influence disease presentation and sample composition. This will affect their representation of the natural disease process. Humane endpoints in animal infection studies may exclude disease stages that occur under field conditions. Each species is unique in how it responds to capture and confinement, as well as sedation and anesthesia regimens and responses to handling by its caretakers. These species-specific characteristics need to be built into study designs to accurately generate representative data based on high-quality samples.

Methods to improve sample quality for molecular-based assays

Specimen quality depends on timely acquisition, adequate cold chain maintenance, and target stabilization. All these requirements are difficult to manage with wildlife studies. To address the needs for stabilizing and transporting biological samples, various studies have been performed, including the use of chemically treated filter paper cards (e.g., Whatman FTA cards; Flinders Technology Associates, GE Healthcare Bio-Sciences, Pittsburgh, PA).^{74,93} Samples collected onto cards are chemically stabilized, do not require immediate refrigeration or freezing, and the cards are lightweight and easily transported.^{22,75,87,101}

Nucleic acids suitable for PCR testing have been recovered from wild mammals, including herpesvirus DNA from Asian elephants (*Elephas maximus*),⁶² DNA from trypanosomes from non-human primates² and African zebu (*Bos taurus indicus*),²⁰ RNA from rabies virus isolates,⁸¹ and DNA for PCR and restriction-fragment length polymorphism (RFLP)-based genetic profiling of wildlife in general.¹⁰¹ In addition to specific research involving sampling from wildlife, the literature includes studies demonstrating stabilization of RNA and DNA from a range of pathogens that can infect wildlife.^{60,61} The quality of a sample preserved on FTA cards has been addressed in studies focused on genetic sequencing of the collected sample. Diagnostic and surveillance testing that require sequence analysis of the sample are more prone to failure given that quality (i.e., stabilization of the genetic materials, requirement for large segments of intact genome) in the sample diminishes compared to testing by conventional PCR and real-time PCR, which are more tolerant of shorter nucleic acid fragments in the test sample.¹¹⁵ Multiple reports have demonstrated effective use of samples collected onto FTA cards for sequence analysis, including genome sequencing of opossums, rodents, and bats⁹; microbiome sequencing from feces of spider monkeys (*Ateles geoffroyi*)⁴² and dogs¹⁰²; genotyping of rotaviruses¹⁰⁹; and automated high-throughput genotyping of cattle using whole blood and nasal samples.⁶⁹

Statistical perspectives: use of LCMs for analysis of test accuracy data from wild mammals of unknown infection status

Bayesian LCMs are appealing to wildlife health specialists because they can be used to evaluate test accuracy in naturally occurring infectious diseases, the sample sizes are often small, and the use of non-invasive methods is greatly preferred for testing species of conservation concern especially when it is prohibited or perhaps unethical to autopsy and collect specimens to establish a definitive diagnosis. Inherent in a Bayesian analysis are the underlying model assumptions of conditional independence of tests,³³ constant sensitivity and specificity across populations,³⁸ and distinct prevalence when applied to scenarios with at least 2 tests and at least 2 populations⁴⁷ and further developed in a Bayesian framework.⁵⁰ For the LCM with 3 tests in a single population, only the assumption of conditional independence is needed for model identifiability. This assumption is most likely correct when the assays target different analytes (e.g., virus isolation, PCR, and serology). A crucial aspect of any Bayesian analysis is proper justification of the prior distributions provided by subject-matter experts. Beta priors, which are commonly used for DSe and DSp, are based on expert guesses of the most likely [modal] value (e.g., DSp = 0.99) and a lower limit (e.g., DSp = 0.95) that the expert is 95% or 99% sure that the true value exceeds. Use of priors centered on the correct parameter values, but which are diffuse enough to not be overly influential, is especially important in non-identifiable models⁵² because prior information about parameters may substantially influence posterior inferences about DSe and DSp. A sensitivity analysis using different priors is considered an essential prerequisite for reporting of a BLCM.⁵⁶ Construction of biologically reasonable Bayesian models that are also statistically sound requires input from biostatisticians, epidemiologists, disease specialists, laboratory diagnosticians, and others who use the tests for clinical decision-making, routine diagnosis, disease surveillance, or research. Bayesian LCMs are readily implemented in freeware platforms such as OpenBUGS⁶⁶ and R (<https://www.r-project.org/>), and their use is facilitated because authors of many studies share code through websites (e.g., <https://cadms.vetmed.ucdavis.edu/diagnostic-tests>).

For study planning purposes, the most common design is to have 2 geographically distinct and unconnected populations in which 2 tests are applied to all sampled animals. Sample sizes depend primarily on estimated disease prevalence in respective populations, whether prior information is available for model parameters, and how close DSe and DSp are to 100%. These factors are described in other papers.^{36,51} In many situations with wild mammal studies, logistical issues constrain the maximum sample size that is obtainable.

Recommendations for future test validation studies and interpretation of test results

Design including statistical analysis

There is no single set of design guidelines for test accuracy studies in wild mammals, but purpose and case definition usually define whether the study design will be based on samples and test results from naturally occurring disease or infection events in populations, samples from animals of known infection status (infected or not), samples from experimental challenge studies, or combinations of 2 or more sources. Each sample type has advantages and disadvantages, and the design needs to balance the potential actions of wildlife health management and practical issues, such as costs and logistics.^{32,49}

For chronic and progressive infections in wild mammals, several components should be considered specifically at the initial planning stage: pathogenesis of the infectious disease in the targeted population, analytical principles of the tests under evaluation, alternative tests and their strengths and limitations, and context (e.g., target population) in which the validation data will be used.⁷⁷ Internal validity (freedom from bias) and external validity (generalizability) of study results are important methodologic qualities that warrant careful consideration by test users.¹¹⁹

The design of the study informs choice of the method of statistical analysis of test results. Guidance on the latter is available elsewhere,¹²⁴ and additional information on LCMs is included earlier in our study and elsewhere.^{36,51} Statistical advice, including the calculation of appropriate sample size, is most useful if done during the planning phase rather than after the data are collected. Although DSe and DSp are the most common metrics of test accuracy, consideration can also be given to alternative measures (e.g., likelihood ratios) and cutoff-independent measures of accuracy (e.g., area under the ROC curve).

Reporting

The reporting of test accuracy studies in wild mammals should follow STARD guidelines¹⁰ given that the only modifications for animals are Standards for Reporting of Animal Diagnostic Accuracy Studies (STRADAS)-aquatic for aquatic animals³⁵ and STRADAS-paratuberculosis,³⁴ which is targeted to a single disease of wild and captive ruminants. STARD assumes that there is a near-perfect reference standard (based on one or more tests) to which results of the test(s) under evaluation are compared. As previously indicated, the assumption of a perfect reference standard is rarely correct, and hence, LCMs are appropriate methods of analysis for many test accuracy studies. STARD-BLCM reporting standards have been published to address key reporting issues in diagnostic accuracy studies that are specific to

Bayesian analysis regardless of the species in which tests were evaluated.⁵⁶

Interpretation

We make the following recommendations about the interpretation of laboratory test results from validation studies and application of the tests in future studies:

1. Have clearly defined questions (purposes and case definition), use the appropriate tests to answer them, and be aware of their limitations. For example, the best serologic tests would be inappropriate to determine the cause of mortality but may be very appropriate for evaluation of prior exposure to pathogens.
2. Clearly define how tests will be interpreted and justify any adjustments. For example, it may be valid to reduce a minimum positive titer threshold for a serologic test if sampling a population well after a disease event or if the intention is to determine the longevity of the antibody response. Analyzing the distribution of data and a graphical display in a histogram or dot diagram may be helpful to compare patterns of recent, mid- and long-term infection in populations.
3. Use multiple testing approaches. For example, detection of BTV and EHDV serotypes can be much more reliable if supported by spatially and temporally matched virologic data even if collected independently of the study. Likewise, the combination of virus isolation, RT-PCR, and next-generation sequencing may provide a broader perspective related to diagnosing a cause of death or clinical outcome.
4. Pay attention to sample quality, develop guidelines, and adopt easily standardized sample collection and preservation technologies, such as FTA cards. Quality control guidelines may be subjective, such as discarding samples with obvious hemolysis or contamination, or could be very specific, such as maintaining a well-defined cold chain. In either case, these considerations are vital to the diagnostic process and should be included in a standard operating procedure. It also may be of value to future researchers to provide some data and analyses supporting potential guidelines for quality assurance.
5. The study population and how it was sampled needs to be defined in detail. This is especially true if seasonal transmission patterns are not considered.
6. Validate and scrutinize laboratory tests, if possible. Simple experimental infections may be possible, or multiple samples with known positive or negative status may be available for inclusion as controls. It also may be helpful to verify results with additional repeat testing and to validate testing using different test platforms, by sequencing, or by testing through independent laboratories.

7. Be cautious with all interpretations. Low detection frequencies of positive results especially should be approached with caution and should be interpreted with the knowledge that positive predictive values will be much reduced at low prevalence. Additionally, DSp estimates do not account for unknown infective agents to which a wild animal may be exposed. In North America, for example, there are numerous exotic BTV and EHDV that may not be currently represented in existing test formats such as virus neutralization, which would lead to false-negative results.⁹²

Acknowledgments

The use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. government.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding


Our study was funded in part through the OIE Collaborating Centre for Diagnostic Test Validation Science in the Asia-Pacific Region.

Supplementary material

Supplementary material for this article is available online.

ORCID iDs

Beibei Jia  <https://orcid.org/0000-0002-8241-4738>

Beate Crossley  <https://orcid.org/0000-0003-2932-7229>

References

1. Al-Adhami BH, Gajadhar AA. A new multi-host species indirect ELISA using protein A/G conjugate for detection of anti-*Toxoplasma gondii* IgG antibodies with comparison to ELISA-IgG, agglutination assay and western blot. *Vet Parasitol* 2014;200:66–73.
2. Aysanoa E, et al. Molecular epidemiology of trypanosomatids and *Trypanosoma cruzi* in primates from Peru. *Ecohealth* 2018;14:732–742.
3. Barasona JA, et al. DNA detection reveals *Mycobacterium tuberculosis* complex shedding routes in its wildlife reservoir the Eurasian Wild Boar. *Transbound Emerg Dis* 2017;64:906–915.
4. Bat Conservation International. Consensus statement of the second WNS Emergency Science Strategy Meeting, Austin, Texas, May 27–28, 2009. <http://208.68.105.19/pdfs/whitenoise/ConsensusStatement2009.pdf>
5. Bingham J, et al. Experimental studies of West Nile virus (New York 1999 and Kunjin virus) indicate low susceptibility of an Australian corvid (Little raven, *Corvus mellori*). *Aust Vet J* 2010;88:204–210.
6. Blasche T, et al. Rapid detection of CWD PrP: comparison of tests designed for the detection of BSE or scrapie. *Transbound Emerg Dis* 2012;59:405–415.

7. Blehert DS, et al. Bat white-nose syndrome: an emerging fungal pathogen? *Science* 2009;323:227.
8. Boadella M, et al. Performance of immunochromatographic and ELISA tests for detecting fallow deer infected with *Mycobacterium bovis*. *Prev Vet Med* 2012;104:160–164.
9. Borisenko AV, et al. DNA barcoding in surveys of small mammal communities: a field study in Suriname. *Mol Ecol Resour* 2008;8:471–479.
10. Bossuyt PM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 2003;138:W1–W12.
11. Brivio F, et al. Assessing the impact of capture on wild animals: the case study of chemical immobilisation on Alpine ibex. *PLoS One* 2015;10:1–18.
12. Cappai S, et al. Evaluation of a commercial field test to detect African swine fever. *J Wildl Dis* 2017;53:602–606.
13. Casais R, et al. Evaluation of an ELISA using recombinant Ss λ 20 Δ B3 antigen for the serological diagnosis of *Sarcoptes scabiei* infestation in domestic and wild rabbits. *Vet Parasitol* 2015;214:315–321.
14. Chambers MA. Review of the diagnosis of tuberculosis in non-bovid wildlife species using immunological methods—an update of published work since 2009. *Transbound Emerg Dis* 2013;60:14–27.
15. Chapinal N, et al. Bayesian analysis to evaluate tests for the detection of *Mycobacterium bovis* infection in free-ranging wild bison (*Bison bison athabascae*) in the absence of a gold standard. *J Wildl Dis* 2015;51:619–625.
16. Chaturvedi S, et al. Rapid real-time PCR assay for culture and tissue identification of *Geomyces destructans*: the etiologic agent of bat geomycosis (white-nose syndrome). *Mycopathologia* 2011;172:247–256.
17. Cohen JF, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016;6:e012799.
18. Colling A, et al. A network approach for provisional assay recognition of a Hendra virus antibody ELISA: test validation with low sample numbers from infected horses. *J Vet Diagn Invest* 2018;30:362–369.
19. Cousins DV, Florisson N. A review of tests available for use in the diagnosis of tuberculosis in non-bovine species. *Rev Sci Tech* 2005;24:1039–1059.
20. Cox AP, et al. Constraints to estimating the prevalence of trypanosome infections in East African zebu cattle. *Parasit Vectors* 2010;3:82.
21. Cuttell L, et al. Evaluation of ELISA coupled with Western blot as a surveillance tool for *Trichinella* infection in wild boar (*Sus scrofa*). *Vet Parasitol* 2014;199:179–190.
22. da Cunha Santos G, et al. FTA cards for preservation of nucleic acids for molecular assays: a review on the use of cytologic/tissue samples. *Arch Pathol Lab Med* 2018;142:308–312.
23. de la Cruz ML, et al. Assessment of the sensitivity and specificity of serological (IFAT) and molecular (direct-PCR) techniques for diagnosis of leishmaniasis in lagomorphs using a Bayesian approach. *Vet Med Sci* 2016;2:211–220.
24. Dickens MJ, et al. Stress and translocation: alterations in the stress physiology of translocated birds. *Proc Biol Sci* 2009;276:2051–2056.
25. Dohoo I, et al. *Veterinary Epidemiologic Research*. 2nd ed. Charlottetown, Canada: AVC, 2010. pp 33–56.
26. Drewe JA, et al. Accuracy of three diagnostic tests for determining *Mycobacterium bovis* infection status in live-sampled wild meerkats (*Suricata suricatta*). *J Vet Diagn Invest* 2009;21:31–39.
27. Drewe JA, et al. Diagnostic accuracy and optimal use of three tests for tuberculosis in live badgers. *PLoS One* 2010;5:e11196.
28. Forzan M, et al. Competitive enzyme-linked immunosorbent assay using baculovirus-expressed VP7 for detection of epizootic haemorrhagic disease virus (EHDV) antibodies. *J Virol Methods* 2017;248:212–216.
29. Gabriel MW, et al. Effectiveness of rapid diagnostic tests to assess pathogens of fishers (*Martes pennanti*) and gray foxes (*Urocyon cinereoargenteus*). *BioOne* 2010;46:966–970.
30. Gall D, et al. Enzyme immunoassays for serological diagnosis of bovine brucellosis: a trial in Latin America. *Clin Diagn Lab Immunol* 1998;5:654–661.
31. Gallardo C, et al. Assessment of African swine fever diagnostic techniques as a response to the epidemic outbreaks in eastern European Union countries: how to improve surveillance and control programs. *J Clin Microbiol* 2015;53:2555–2565.
32. Gardner IA, et al. Validity of using serological tests for diagnosis of diseases in wild animals. *Rev Sci Tech* 1996;15:323–335.
33. Gardner IA, et al. Conditional dependence between tests affects the diagnosis and surveillance of animal diseases. *Prev Vet Med* 2000;45:107–122.
34. Gardner IA, et al. Consensus-based reporting standards for diagnostic test accuracy studies for paratuberculosis in ruminants. *Prev Vet Med* 2011;101:18–34.
35. Gardner IA, et al. Recommended reporting standards for test accuracy studies of infectious diseases of finfish, amphibians, molluscs and crustaceans: the STRADAS-Aquatic checklist. *Dis Aquat Organ* 2016;118:91–111.
36. Gardner IA, et al. Design, statistical analysis and reporting standards for test accuracy studies for infectious diseases in animals: progress, challenges and recommendations. *Prev Vet Med* 2019; 62:46–55.
37. Gargas A, et al. *Geomyces destructans* sp. nov. associated with bat white-nose syndrome. *Mycotaxon* 2009;108:147–154.
38. Gorsich EE, et al. Evaluation of the sensitivity and specificity of an enzyme-linked immunosorbent assay for diagnosing brucellosis in African buffalo (*Syncerus caffer*). *J Wildl Dis* 2015;51:9–18.
39. Gowtage-Sequeira S, et al. Evaluation of the CervidTB STAT-PAK for the detection of *Mycobacterium bovis* infection in wild deer in Great Britain. *Clin Vaccine Immunol* 2009;16:1449–1452.
40. Greiner M, Gardner IA. Epidemiologic issues in the validation of veterinary diagnostic tests. *Prev Vet Med* 2000;45:3–22.
41. Haas C, et al. Evaluation of a commercial ELISA for the detection of antibodies to *Sarcoptes scabiei* in wild boar (*Sus scrofa*). *J Wildl Dis* 2015;51:729–733.
42. Hale V, et al. Effect of preservation method on spider monkey (*Ateles geoffroyi*) fecal microbiota over 8 weeks. *J Microbiol Methods* 2015;113:16–26.
43. Halpin K, et al. Pteropid bats are confirmed as the reservoir hosts of henipaviruses: a comprehensive experimental study of virus transmission. *Am J Trop Med Hyg* 2011;85:946–951.
44. Hanon JB, et al. Inter-laboratory evaluation of the performance parameters of a lateral flow test device for the detection of bluetongue virus-specific antibodies. *J Virol Methods* 2016;228:140–150.

45. Himsworth CG, et al. Comparison of test performance and evaluation of novel immunoassays for tuberculosis in a captive herd of wood bison naturally infected with *Mycobacterium bovis*. *J Wildl Dis* 2010;46:78–86.
46. Howerth EW. Bluetongue, epizootic hemorrhagic disease, and other orbivirus-related diseases. In: Williams ES, Barker IK, eds. *Infectious Diseases of Wild Mammals*. Ames, IA: Iowa State University Press, 2001:77–97.
47. Hui SL, Walter SD. Estimating the error rates of diagnostic tests. *Biometrics* 1980;36:167–171.
48. Inagaki E, et al. Development of a serodiagnostic multi-species ELISA against tick-borne encephalitis virus using subviral particles. *Ticks Tick Borne Dis* 2016;7:723–729.
49. Jacobson RH. Validation of serological assays for diagnosis of infectious diseases. *Rev Sci Tech* 1998;17:469–526.
50. Johnson WO. Screening without a “gold standard”: the Hui-Walter paradigm revisited. *Am J Epidemiol* 2001;153:921–924.
51. Johnson WO, et al. On the interpretation of test sensitivity in the two-test two-population problem: assumptions matter. *Prev Vet Med* 2009;91:116–121.
52. Jones G, et al. Identifiability of models for multiple diagnostic testing in the absence of a gold standard. *Biometrics* 2010;66:855–863.
53. Katz J, et al. Diagnostic analysis of the prolonged bluetongue virus RNA presence found in the blood of naturally infected cattle and experimentally infected sheep. *J Vet Diagn Invest* 1994;6:139–142.
54. Keet DF, et al. Intradermal tuberculin testing of wild African lions (*Panthera leo*) naturally exposed to infection with *Mycobacterium bovis*. *Vet Microbiol* 2010;144:384–391.
55. Kienzle C, et al. Virus isolation and molecular detection of bluetongue and epizootic hemorrhagic disease viruses from naturally infected white-tailed deer (*Odocoileus virginianus*). *J Wildl Dis* 2017;53:843–849.
56. Kostoulas P, et al. STARD-BLCM: Standards for the Reporting of Diagnostic accuracy studies that use Bayesian Latent Class Models. *Prev Vet Med* 2017;138:37–47.
57. Laurin E, et al. Design standards for experimental and field studies to evaluate diagnostic accuracy of tests for infectious diseases in aquatic animals. *J Fish Dis* 2018;41:729–749.
58. Le Dimna M, et al. Validation of two commercial real-time RT-PCR kits for rapid and specific diagnosis of classical swine fever virus. *J Virol Methods* 2008;147:136–142.
59. Levavasseur E, et al. Detection and partial discrimination of atypical and classical bovine spongiform encephalopathies in cattle and primates using real-time quaking-induced conversion assay. *PLoS One* 2017;12:1–15.
60. Liang X, et al. Evaluation of fast technology analysis (FTA) cards as an improved method for specimen collection and shipment targeting viruses associated with bovine respiratory disease complex. *J Virol Methods* 2014;202:69–72.
61. Linhares DCL, et al. Evaluation of Flinders Technology Associates cards for collection and transport of samples for detection of porcine reproductive and respiratory syndrome virus by reverse transcription polymerase chain reaction. *J Vet Diagn Invest* 2012;24:328–332.
62. Lopez J, et al. Assessment of a Lance-and-Swab blood sampling technique for surveillance of elephant endotheliotropic herpesvirus infection. *J Zoo Wildl Med* 2017;48:659–667.
63. Lorch JM, et al. Rapid polymerase chain reaction diagnosis of white-nose syndrome in bats. *J Vet Diagn Invest* 2010;22:224–230.
64. Lorch JM, et al. Experimental infection of bats with *Geomyces destructans* causes white-nose syndrome. *Nature* 2011;480:376–378.
65. Lorch JM, et al. A culture-based survey of fungi in soil from bat hibernacula in the eastern United States and its implications for detection of *Geomyces destructans*, the causal agent of bat white-nose syndrome. *Mycologia* 2013;105:237–252.
66. Lunn D, et al. The BUGS project: evolution, critique and future directions. *Stat Med* 2009;28:3049–3067.
67. Lyashchenko KP, et al. Animal-side serologic assay for rapid detection of *Mycobacterium bovis* infection in multiple species of free-ranging wildlife. *Vet Microbiol* 2008;132:283–292.
68. Lyashchenko KP, et al. Rapid detection of serum antibody by dual-path platform VetTB assay in white-tailed deer infected with *Mycobacterium bovis*. *Clin Vaccine Immunol* 2013;20:907–911.
69. McClure MC, et al. Assessment of DNA extracted from FTA cards for use on the Illumina iSelect BeadChip. *BMC Res Notes* 2009;2:107.
70. McInnes MDF, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *J Am Med Assoc* 2018;319:388–396.
71. Meteyer CU, et al. Histopathologic criteria for confirming white-nose syndrome in bats. *J Vet Diagn Invest* 2009;21:411–414.
72. Minnis AM, Lindner DL. Phylogenetic evaluation of *Geomyces* and allies reveals no close relatives of *Pseudogymnoascus destructans*, comb. nov., in bat hibernacula of eastern North America. *Fungal Biol* 2013;117:638–649.
73. Muller LK, et al. Bat white-nose syndrome: a real-time TaqMan polymerase chain reaction test targeting the intergenic spacer region of *Geomyces destructans*. *Mycologia* 2013;105:253–259.
74. Muthukrishnan M, et al. Evaluation of FTA cards as a laboratory and field sampling device for the detection of foot-and-mouth disease virus and serotyping by RT-PCR and real-time RT-PCR. *J Virol Methods* 2008;151:311–316.
75. Natarajan P, et al. Paper-based archiving of mammalian and plant samples for RNA analysis. *Biotechniques* 2000;29:1328–1333.
76. Nelson JT, et al. Evaluation of serodiagnostic assays for *Mycobacterium bovis* infection in elk, white-tailed deer, and reindeer in the United States. *Vet Med Int* 2012;2012:563293.
77. Nielsen SS, et al. Structured approach to design of diagnostic test evaluation studies for chronic progressive infections in animals. *Vet Microbiol* 2011;150:115–125.
78. O’Brien DJ, et al. Evaluation of blood assays for detection of *Mycobacterium bovis* in white-tailed deer (*Odocoileus virginianus*) in Michigan. *J Wildl Dis* 2009;45:153–164.
79. O’Brien DJ, et al. Performance of diagnostic tests for bovine tuberculosis in North American furbearers and implications for surveillance. *Transbound Emerg Dis* 2013;60:67–73.
80. Peel AJ, et al. Use of cross-reactive serological assays for detecting novel pathogens in wildlife: assessing an appropriate cutoff for henipavirus assays in African bats. *J Virol Methods* 2013;193:295–303.

81. Picard-Meyer E, et al. Use of filter paper (FTA) technology for sampling, recovery and molecular characterization of rabies viruses. *J Virol Methods* 2007;140:174–182.
82. Pollock JM, et al. Immune responses in bovine tuberculosis. *Tuberculosis* 2001;81:103–107.
83. Prieto JM, et al. Sensitive and specific enzyme-linked immunosorbent assay for detecting serum antibodies against *Mycobacterium avium* subsp. *paratuberculosis* in fallow deer. *Clin Vaccine Immunol* 2014;21:1077–1085.
84. Pruvot M, et al. The modification and evaluation of an ELISA test for the surveillance of *Mycobacterium avium* subsp. *paratuberculosis* infection in wild ruminants. *BMC Vet Res* 2013;9:1–8.
85. Quist CF, et al. Host defense responses associated with experimental hemorrhagic disease in white-tailed deer. *J Wildl Dis* 1997;33:584–599.
86. Ráez-Bravo A, et al. Evaluation of three enzyme-linked immunosorbent assays for sarcoptic mange diagnosis and assessment in the Iberian ibex, *Capra pyrenaica*. *Parasit Vectors* 2016;9:1–8.
87. Rodriguez-Prieto V, et al. Porcine reproductive and respiratory syndrome (PRRS) virus in wild boar and Iberian pigs in south-central Spain. *Eur J Wildl Res* 2013;59:859–867.
88. Roelke ME, et al. T-lymphocyte profiles in FIV-infected wild lions and pumas reveal CD4 depletion. *J Wildl Dis* 2006;42:234–248.
89. Romero LM. Physiological stress in ecology: lessons from biomedical research. *Trends Ecol Evol* 2004;19:249–255.
90. Roos EO. Test performance of three serological assays for the detection of *Mycobacterium bovis* infection in common warthogs (*Phacochoerus africanus*). *Vet Immunol Immunopathol* 2016;182:79–84.
91. Ruder MG, et al. The first 10 years (2006–15) of epizootic hemorrhagic disease virus serotype 6 in the USA. *J Wildl Dis* 2017;53:901–905.
92. Ruder MG, et al. Transmission and epidemiology of bluetongue and epizootic hemorrhagic disease in North America: current perspectives, research gaps, and future directions. *Vector Borne Zoonotic Dis* 2015;15:348–363.
93. Sakai T, et al. Establishing conditions for the storage and elution of rabies virus RNA using FTA cards. *J Vet Med Sci* 2015;77:461–465.
94. Sastre P, et al. Development of a novel lateral flow assay for detection of African swine fever in blood. *BMC Vet Res* 2016;12:206.
95. Schmidt RL, Factor RE. Understanding sources of bias in diagnostic accuracy studies. *Arch Pathol Lab Med* 2013;137:558–565.
96. Schumaker BA, et al. Evaluation of the western immunoblot as a detection method for *Brucella abortus* exposure in elk. *J Wildl Dis* 2010;46:87–94.
97. Scoles GA, et al. Validation of an *Anaplasma marginale* cELISA for use in the diagnosis of *A. ovis* infections in domestic sheep and *Anaplasma* spp. in wild ungulates. *Vet Microbiol* 2008;130:184–190.
98. Shuey MM, et al. A highly sensitive qPCR assay for the detection and differentiation of *Pseudogymnoascus destructans* and *Pseudogymnoascus* species. *Appl Environ Microbiol* 2014;80:1726–1731.
99. Shury TK, et al. Field evaluation of three blood-based assays for elk (*Cervus canadensis*) naturally infected with *Mycobacterium bovis*. *Prev Vet Med* 2014;115:109–121.
100. Skerratt LF, et al. Validation of diagnostic tests in wildlife: the case of chytridiomycosis in wild amphibians. *J Herpetol* 2011;45:444–450.
101. Smith LM, Burgoyne LA. Collecting, archiving and processing DNA from wildlife samples using FTA databasing paper. *BMC Ecol* 2004;4:4.
102. Song SJ, et al. Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. *mSystems* 2016;1:e00021-16.
103. Stallknecht DE, et al. Epizootic hemorrhagic disease virus and bluetongue virus serotype distribution in white-tailed deer in Georgia. *J Wildl Dis* 1995;31:331–338.
104. Stallknecht DE, et al. Hemorrhagic disease in white-tailed deer in Texas: a case for enzootic stability. *J Wildl Dis* 1996;32:695–700.
105. Stallknecht DE. Impediments to wildlife disease surveillance, research, and diagnostics. *Curr Top Microbiol Immunol* 2007;315:445–461.
106. Stephenson N, et al. Development and validation of a fecal PCR assay for *Notoedres cati* and application to notoedric mange cases in bobcats (*Lynx rufus*) in northern California, USA. *J Wildl Dis* 2013;49:303–311.
107. Stewart LD, et al. Development of a novel immunochromatographic lateral flow assay specific for *Mycobacterium bovis* cells and its application in combination with immunomagnetic separation to test badger faeces. *BMC Vet Res* 2017;13:1–12.
108. Sylvester TT, et al. Prevalence and risk factors for *Mycobacterium bovis* infection in African lions (*Panthera leo*) in the Kruger National Park. *J Wildl Dis* 2017;53:372–376.
109. Tam KI, et al. Evaluation of BBL Sensi-Discs and FTA cards as sampling devices for detection of rotavirus in stool samples. *J Virol Methods* 2015;222:41–46.
110. Tankaew P, et al. Evaluation of an in-house indirect ELISA for detection of antibody against haemorrhagic septicemia in Asian elephants. *J Microbiol Methods* 2017;134:30–34.
111. Thomsen BV, et al. Diagnostic accuracy of rectal mucosa biopsy testing for chronic wasting disease within white-tailed deer (*Odocoileus virginianus*) herds in North America: effects of age, sex, polymorphism at PRNP codon 96, and disease progression. *J Vet Diagn Invest* 2012;24:878–887.
112. Tignon M, et al. Development and inter-laboratory validation study of an improved new real-time PCR assay with internal control for detection and laboratory diagnosis of African swine fever virus. *J Virol Methods* 2011;178:161–170.
113. U.S. Geological Survey (USGS). Revised case definitions for white-nose syndrome and winter 2010/2011 bat submission guidelines. NWHC Wildlife Health Bulletin 2011-01. Available from: <https://www.usgs.gov/media/files/revised-case-definitions-white-nose-syndrome>.
114. van der Heijden EMDL, et al. Field application of immunoassays for the detection of *Mycobacterium bovis* infection in the African buffalo (*Syncerus caffer*). *Vet Immunol Immunopathol* 2016;169:68–73.
115. Vilcek S, et al. The first detection of PRRSV in wild boars in Slovakia. *Berl Munch Tierarztl Wochenschr* 2015;128:31–33.

116. Wadhwa A, et al. Evaluation of ethanol vortex ELISA for detection of bovine tuberculosis in cattle and deer. *BMC Vet Res* 2014;10:147.
117. Wallander C, et al. *Toxoplasma gondii* seroprevalence in wild boars (*Sus scrofa*) in Sweden and evaluation of ELISA test performance. *Epidemiol Infect* 2015;143:1913–1921.
118. Wasniewski M, et al. Evaluation of an ELISA to detect rabies antibodies in orally vaccinated foxes and raccoon dogs sampled in the field. *J Virol Methods* 2013;187:264–270.
119. Whiting P, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
120. Wobeser GA. How disease is detected, described, and measured. In: *Essentials of Disease in Wild Animals*. Hoboken, NJ: Wiley Blackwell, 2006:45–60.
121. World Organisation for Animal Health (OIE). Principles and methods of validation of diagnostic assays for infectious diseases, Chapter 1.1.6. In: *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. Paris, France: OIE, 2018.
122. World Organisation for Animal Health (OIE). Principles and methods of validation of diagnostic tests for infectious disease applicable to wildlife, Chapter 2.2.7. In: *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. Paris, France: OIE, 2018.
123. World Organisation for Animal Health (OIE). Bluetongue (infection with bluetongue virus), Chapter 3.1.3. In: *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. Paris, France: OIE, 2018.
124. World Organisation for Animal Health (OIE). Statistical approaches to validation, Chapter 2.2.5. In: *Manual of Diagnostic Tests and Vaccines for Terrestrial Animals*. Paris, France: OIE, 2018.
125. Yabsley MJ, et al. Evaluation of the Chagas Stat-Pak assay for detection of *Trypanosoma cruzi* antibodies in wildlife reservoirs. *J Parasitol* 2009;95:775–777.
126. Yang DK, et al. Comparison of four diagnostic methods for detecting rabies viruses circulating in Korea. *J Vet Sci* 2012;13: 43–48.