Methods in Ecology and Evolution

ERLEND B. BIRKELAND NILSEN (Orcid ID : 0000-0002-5119-8331)

DR ATLE MYSTERUD (Orcid ID : 0000-0001-8993-7382)

Article type : Research Article

Editor : Professor Nigel Yoccoz

A method that accounts for differential detectability in mixed samples of long-term infections with applications to the case of Chronic Wasting Disease in cervids

Hildegunn Viljugrein[1], Petter Hopp[1], Sylvie L. Benestad[1], Erlend B. Nilsen[2], Jørn Våge[1], Saraya Tavornpanich[1], Christer M. Rolandsen[2], Olav Strand[2] and Atle Mysterud[3*]

[1] Norwegian Veterinary Institute, P.O. Box 750 Sentrum, NO-0106 Oslo, Norway.

[2] Norwegian Institute for Nature Research (NINA), P. O. Box 5685 Torgarden, NO-7485 Trondheim, Norway.

[3] Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, P.O. Box 1066 Blindern, NO-0316 Oslo, Norway.

[*] Corresponding author: atle.mysterud@ibv.uio.no

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/2041-210X.13088

Summary

1. Surveillance of wildlife diseases is logistically difficult, and imperfect detection is a recurrent challenge for disease estimation. Using citizen science can increase sample sizes, but it is associated with a cost in terms of the anatomical type and quality of the sample. Additionally, biological tissue samples from remote areas lose quality due to autolysis. These challenges are faced in the case of emerging Chronic Wasting Disease (CWD) in cervids.

2. Here, we develop a stochastic scenario tree model of diagnostic sensitivity, allowing for a mixture of tissue sample types (lymph nodes and brain) and qualities while accounting for different detection probabilities during the CWD infection, lasting 2-3 years. We apply the diagnostic sensitivity in a Bayesian framework, enabling estimation of age-class-specific true prevalence, including the prevalence in latent, recently infected stages. We provide a simulation framework to estimate the sensitivity of the surveillance system (i.e., the probability of detecting the infection in a given population), when detectability varies among individuals due to different disease progression.

3. We demonstrate the utility of our framework by applying it to the recent emergence of CWD in a European population of reindeer. We estimated apparent CWD prevalence at 1.2 % of adults in the infected population of wild reindeer, while the true prevalence was 1.6 %. The sensitivity estimation of the CWD surveillance was performed in an adjacent small (~500) and a large (~10,000) reindeer population, demonstrating low certainty of CWD absence.

4. Our method has immediate application to the mandatory testing for CWD in EU countries commencing in 2018. Similar approaches that account for latent stages and a serial disease progression in various tissues with a temporal pattern of diagnostic sensitivity may enhance the estimation of the prevalence of wildlife diseases more generally.

Key-words: Bayesian estimation methods, surveillance, wildlife diseases, diagnostic sensitivity, test sensitivity, chronic wasting disease, prevalence, prions, PrP^{CWD}

Introduction

Management of emerging wildlife diseases is a challenge in many parts of the world (Jones et al., 2008). Mitigating diseases of wildlife often requires estimation of prevalence at a population level, either to determine status or whether a disease is emerging or not (Funk et al., 2013). Equally important is the assessment of the probability of detecting a disease in a population for a certain design prevalence, i.e., the sensitivity of the surveillance system (Martin, Cameron & Greiner 2007). In early epidemic stages, the prevalence is low and large sample sizes are required to detect the infection when using active surveillance (Doherr & Audigé 2001). Therefore, also establishing the absence of disease requires large efforts. For wildlife populations, sample sizes are often limited due to logistical constraints, and samples may be of various quality due to hunters' involvement in sampling (Rhyan & Spraker 2010). Tackling such uncertainty is important for estimation processes but is rarely accounted for explicitly.

Chronic wasting disease (CWD) in cervids is a transmissible spongiform encephalopathy involving prions as the infectious agent (Williams & Young 1980). CWD was first recognized in 1967 in captive mule deer (*Odocoileus hemionus*) in Colorado, USA, and emerged in wild deer in 1981 (Miller et al., 2000). It has now spread to 25 states in the USA and 2 provinces in Canada and to South Korea after importing infected stock from Canada. The first case of CWD in Europe was detected in a female reindeer (*Rangifer tarandus*) in March 2016 in the Nordfjella mountain range of Norway (Benestad et al., 2016). Since then, sampling by aid of hunters and recently, governmental culling, has revealed 19 cases in total in this population, while no CWD-positive cases have been found in adjacent populations. However, it is not clear whether this represents a limited spread of disease or just a lack of detection, as evidence from USA sources shows that the prevalence levels of CWD are low for

years after the introduction into a population (Heisey et al., 2010). From a management perspective, it is urgent to establish methods to assess the CWD prevalence of infected populations and, even more importantly, to estimate the probability of detecting CWD in populations adjacent to an infected population. The CWD-infected population in Norway has limited, but not absent, physical connectivity with adjacent populations. Furthermore, adjacent populations are candidates for repopulating Nordfjella zone 1 after a period of fallowing. Therefore, the ability to document freedom from CWD at a defined design prevalence is critical.

A key feature of CWD is the long incubation period (Davenport et al., 2018). The time from infection to death is in the range of 1.5-2.5 years in mule deer (Fox et al., 2006). Due to low initial detectability and slow disease progression, it is rare to detect infection in juveniles, and the prevalence is lower in yearlings than in adults (Samuel & Storm 2016). Prions causing CWD (PrP^{CWD}) most often enter the animal through the oronasal cavity and move into the body via the lymphoid system, while it often takes a year before brain tissue is affected (Davenport et al., 2017; Hoover et al., 2017). Therefore, the type of tissue used will affect the sensitivity of the diagnostic tests for CWD (Keane et al., 2008). Retropharyngeal lymph nodes are the preferred tissue enabling early detection but may be difficult to identify for non-trained personnel. Due to logistical constraints and the combined use of trained professionals and hunters to collect samples of reindeer in the remote mountain ranges of Norway, a mixture of samples was collected. Samples include either brain tissue only or both brain and lymph nodes, making coherent estimation of prevalence and detection of disease challenging with a complex error structure.

Here, we develop a simulation (scenario tree) model to estimate the test sensitivity of CWD with a mixture of samples affecting detectability in early (retropharyngeal lymph nodes), middle (obex) and late (brainstem) stages during the disease development (Fig. 1). We develop two applications. Application 1 is an estimation of age- and infection-stage-specific true prevalence within a Bayesian framework, applied here to the only known CWD-infected population of reindeer in Norway.

Application 2 is a framework for estimation of the surveillance sensitivity in a given population with unknown CWD status, applied here to adjacent wild reindeer populations differing in size, sample type and quality.

Material and methods

Study areas and the population estimation model

Our study areas were the Nordfjella and Hardangervidda high alpine mountain ranges of Norway. CWD was detected in management zone 1 of Nordfjella (2000 km^2) with ~2000 reindeer. Surveillance was also initiated in the adjacent herds in Nordfjella management zone 2 (1000 km^2) with ~500 reindeer and on Hardangervidda (8000 km^2) with ~10,000 reindeer. We estimated the reindeer population size and age and sex structure in each area using a hierarchical change-in-ratio model with parameters estimated using Bayesian inference, based on aerial minimum counts in winter, calving surveys in summer, harvest data and demographic structure counts in fall. Further details of study areas, census data and the population estimation model are shown in Appendix S1.

Sampling of lymph nodes and brain tissue

Due to logistical and economic constraints, sampling regimes differed for the smaller and more accessible Nordfjella area compared to the larger and more remote Hardangervidda area. In Nordfjella, hunters delivered whole heads to six stations in zone 1 and two stations in zone 2. Heads collected were then transported to the two most centrally located stations where local trained veterinarians sampled the brainstem (*medulla oblongata*) and the retropharyngeal lymph nodes (RLN). In rare cases, the RLN had unintentionally been removed by hunters and alternative head lymph nodes (submandibular, parotid) were collected. On Hardangervidda, hunters were provided with a sampling kit including gloves, a sampling spoon to harvest brainstem, a tissue container and a prepaid postal envelope.

Model of disease progression

A main feature of our model is the incorporation of how CWD infection stages affect detection probability in different tissues (Fig. 1). The RLN are typically infected before brain tissue (stage 1). The primary site of infection in the brain is the dorsal motor of the vagus nerve (DMVN) at the level of the obex part of the brainstem (stage 2), while the whole brainstem gradually becomes infected in later stages (stage 3). The main timeline is drawn for an expected two-year progression from infection to death. Brainstem accumulation of PrP^{CWD}, as assessed with the obex score from immunohistochemistry (IHC) staining, was used to estimate disease progression (Thomsen et al., 2012; Haley & Richt 2017):

- Obex score 0: Not detected.
- Obex score 1: Minimal to moderate IHC staining in the DMNV part of obex only.
- Obex score 2: Moderate to heavy staining in the DMVN and moderate staining in the adjacent nuclei and tracts.
- Obex score 3: Intense staining in the DMVN and solitarious nucleus, moderate staining in other nuclei.
- Obex score 4: Prominent staining in all the nuclei of the obex area and staining visible in the whole brainstem.

For RLN, we used:

- RLN score 0: Not detected.
- RLN score 1: Immunostaining in few follicles, low (<1) optical density (OD) values by ELISA (TeSeE SAP ELISA test from Bio-Rad).
- RLN score 2: Immunostaining in the majority of the follicles, high (>1) OD values by ELISA.

We hence separated the following stages of infection (Fig. 1):

- Stage 0 (RLN-, obex-, brainstem-): The animal has recently acquired the infection and the amount of PrP^{CWD} is below detectable levels in both the RLN and brain tissue. Stage 0 does not include true CWD negative animals.

- Stage 1 (RLN+, obex-, brainstem-): RLN can become positive three months post-infection. The distribution of PrP^{CWD} is not uniformly distributed in RLN. The sampling of RLN tissue may be negative if the few lymphoid follicles containing PrP^{CWD} are missing. The amount of PrP^{CWD} is steadily increasing but levels off and reaches a plateau after ~ 6 months.
- Stage 2 (RLN+, obex+, brainstem-): At approximately 9 months post-infection, PrP^{CWD} are detectable in some structures in the obex area. Sample quality is hence a main issue, as the amount of PrP^{CWD} is minimal outside the obex part of the brainstem.
- Stage 3 (RLN+, obex+, brainstem+): After ~12 months post-infection, PrP^{CWD} is also detectable outside the obex area, in the rest of the *medulla oblongata*. The presence of obex tissue is no longer crucial to detect infection.

As the duration of CWD infection varies also depending on the prion protein gene (*PRNP*) and the prion strain and species involved, we also tested how extending to a 3-year disease progression affected estimation (Fig. 3B). A prolonged disease progression was modelled by stretching the curve(s), making expected death extend to three rather than two years.

Sample quality

Variation in sample quality arises for different reasons. PrP^{CWD} is resistant enough to be detected in autolysed tissues, but it is impossible to assure that the analysed field sample represents obex tissue due to anatomical disintegration. In addition, it is more likely that trained personnel will obtain an appropriate obex sample, while unexperienced hunters may harvest samples lacking anatomical integrity, despite fresh tissues. Hence, early in stage 2, when PrP^{CWD} is detectable only in the obex area, the sample quality is of major importance, diminishing as infection spreads outside the obex area and into the brainstem. Due to logistical constraints in mass testing, it was not possible to record the sample quality. Therefore, the 13 laboratory technicians involved were asked (23rd of October 2017) for their expert opinion, and they reported that 20 % of the samples were without identifiable obex and hence of low quality (Table 1).

Analytical tests

The primary test was an ELISA (TeSeE® ELISA SAP, Bio-Rad, Hercules, CA, USA). A positive or inconclusive result was confirmed by western blot (TeSeE® Western Blot, Bio-Rad, Hercules, CA, USA). The analytical tests have close to perfect specificity (European Food Safety Authority (EFSA) 2005). The analytical test sensitivity for the ELISA is 92.5 % (81.8 – 97.9) for obex and 98.8 % (93.5 – 99.97) for RLN (Hibler et al., 2003), but note that we pooled samples (Appendix S2).

Stochastic scenario trees to model diagnostic test sensitivity

Diagnostic test sensitivity (dSe) is the probability of correctly classifying an infected individual as positive by the diagnostic test. We used stochastic scenario trees (Martin, Cameron & Greiner 2007) to explicitly model the relationship between dSe, sample quality and disease progression (Fig. 2). For each of 10,000 iterations, we randomly drew a point along the axis of time since infection occurred (Fig. 1). Due to the long disease development with sampling in the autumn, the time since infection will be a maximum of 2 years in adults, 17 months in yearlings and 5 months in calves. A given testing regime was specified by the proportion of samples, including RLN (*PrRLN*), the proportion of samples according to age class (*PrAdult, PrYearling*), the proportion of low-quality obex samples (*PrLQ*), and the tissue-specific sensitivity of the ELISA-test according to disease progression of individual *i* ($SeRLN_i$, $SeLQ_i$, $SeHQ_i$). This determines how the simulated infected animals distribute through the scenario tree (Fig. 2). The tissue-specific sensitivity according to the progression of infection was assumed to be a direct function of the expected RLN / obex score at a given time since infection (Table 1). Each simulated individual ends up in one of the limbs of the scenario tree, with a corresponding individual sensitivity according to the sample, time since infection and the respective stochasticity (as specified in Table 1). dSe was modelled as a function of time since infection, by scaling scores above 1 in Figure 1, between the lower and upper level as given by SeLow (Table 1) and the analytical test sensitivity, respectively. The dSe of age class c is then given by the distribution of 10,000 simulated individuals of age class c. The dSe was utilized in two applications: 1) prevalence estimation and 2) estimating surveillance sensitivity.

The proportion of the simulated individuals with no positive outcome was reported as the probability of the non-detectable infected sample (Fig. 3, Appendix S3). The non-zero part of the simulated distribution of the dSe was fitted by a beta distribution, utilizing the R package "fitdistrplus" (Delignette-Muller & Dutang 2015) and the function "fitdist" for each age class and tissue type. The resulting dSe is then age class, tissue and individual specific and consists of two parts (Appendix S3): 1) the distribution of detectable samples (dSe>0) and 2) the probability of non-detectable, early stage (stage 0 with RLN and obex sample; stage 0+1 with obex only sample) samples (Fig. 3C,D).

Application 1: Prevalence estimation

CWD-data from Nordfjella consisted of the number of diagnosed positive and negative per age class. We included animals harvested and tested after the first CWD case was detected in Norway and until the 1st of November 2017 when ordinary hunting ended (Appendix S4). In addition, one positive reindeer sampled from a rectal biopsy during GPS collaring and found positive was included. In total, the data included 2 and 4 positive cases of adults, in 2016 and 2017, respectively. For 2017, information of age class was lacking for 75 of 350 samples. We distributed the unknown samples according to the age-class ratio of the known samples.

Bayesian approaches to prevalence estimation usually assume that the number of positives is approximated by the binomial or Poisson distribution (Lewis & Torgerson 2012). In our case, the population is small compared to the number of individuals tested. Hence, we chose to use the hypergeometric distribution, which accounts for sampling without replacement from a finite population. The prevalence was estimated separately for calves, yearlings and adults (≥2 years). The hypergeometric distribution (dhyper) specifies the number of positive samples (k_c) to be distributed around the number of positives in the population that are likely to be detected ($Ninf_c$ - $N0_c$ - $N1_c$), so that

$k_c \sim$ dhyper ($Ninf_c$ - $N0_c$ - $N1_c$, m_c, $nSample_c$)

where $Ninf_c$ is the number of infected animals (all infection stages 0-3) of age class c, $N0_c$ is the number of early-stage infected animals that are non-detectable (dSe=0), $N1_c$ is the expected number of detectable animals that are not detected due to imperfect test sensitivity, and m_c is the number of non-infected animals in the population plus the expected number of infected animals not being detected ($N0_c$+$N1_c$). Together, ($Ninf_c$ - $N0_c$ - $N1_c$) + m_c constitutes the number of individuals in age class c ($PopSize_c$). The number of individuals tested was given by $nSample_c$. *Apparent prevalence* (AP) is then the proportion of animals from a representative sample of the population that are positive with the diagnostic method used and, specifically, for each age class c: AP_c = ($Ninf_c$ - $N0_c$ - $N1_c$)/$PopSize_c$

Going from AP to *true prevalence* (TP), we utilized the diagnostic sensitivity as modelled above (Fig. 3; Appendix S3). Non-detectable infected animals ($N0_c$) were coming from stage 0, if both the RLN and obex were tested, and as a proportion from stage 1 that was dependent on the proportion of samples tested without RLN (1 – pRLN). Let $p0RLNobex_c$ represent the probability of an infected reindeer of age class c being non-detectable by RLN, and similarly, let $p0obex_c$ represent the probability of being non-detectable by only testing the obex. Then, for each age class c, the probability of the non-detectable stage $Prob\_NonDetectable_c$ = $p0obex_c$ *(1-pRLN) + $p0RLNobex_c$*pRLN. We assume that the number of infected individuals in Nordfjella that are non-detectable ($N0_c$) can be estimated by drawing from a binomial distribution:

$N0_c \sim$ dbin($NInf_c$, prob= $Prob\_NonDetectable_c$)

Similarly, we assume that the number of infected individuals in the early stage 0 (Nr_stage0_c) can be estimated by drawing from a binomial distribution:

$Nr\_stage0_c \sim$ dbin($NInf_c$, prob= $p0RLNobex_c$).

The TP of age class c is the proportion of infected animals (stage 0-3):

$TP_c = Ninf_c / PopSize_c$.

The testing regime that was used provided an expected diagnostic sensitivity (defined by a stochastic distribution) for the detectable part of the infected animals ($Ninf_c - N0_c$): $dSe_c = dSe\_obex_c * (1-pRLN) + dSe\_RLNobex_c * pRLN$, where dSe_obex_c and $dSe\_RLNobex_c$ are specified by the beta distributions in Appendix S3. The expected number of infected individuals that were detectable (dSe>0) but were not detected due to imperfect sensitivity is:

$N1_c = Ninf_c - NrPos_c - N0_c$, where the number of infected being detected as positive was: $NrPos_c \sim dbin(Ninf_c - N0_c, prob = dSe_c)$.

The probability of an infected reindeer of age class *c* being detectable by RLN, but obex negative, is given by *p0obex.RLNpos* (Appendix S3). For the Nordfjella population and according to the simulation model for dSe, we assume that the number of individuals in stage 1 and stage 2+3 can be estimated by drawing from a binomial distribution:

$Nr\_stage1_c \sim dbin(NInf_c, prob= p0obex.RLNpos_c)$

$Nr\_stage23_c \sim dbin(NInf_c, prob= (1-p0obex_c))$

We started by estimating the prevalence in 2016. We set priors for calves and yearlings to be on average 0.10 and 0.36 of the adult prevalence, respectively (Samuel & Storm 2016). We then estimated the prevalence in 2017 by using the estimated TP of 2016 as a prior (mean and standard deviation of 2016-prevalence were used to define beta parameters for the prior-distribution of 2017-prevalence). For calves, there was no information in the TP (wide standard deviation of estimate) of year 2016, so we kept the prevalence-ratio assumption that $TP_{calves} = a1 * TP_{adults}$, where a1 is beta-distributed with a mean of 0.1 (range: 0 - 0.5; Appendix S5).

The prevalence modelling using Bayesian inference was implemented in R through the library R2JAGS (linking R with JAGS version 4.2.0). Prior distributions are required for the unknown model parameters, and Markov chain Monte Carlo simulations are used to find the posterior distribution and to sample from it (Lunn et al., 2013). Mean and median values with 95 % posterior (credible) intervals were calculated as the 0.025 and 0.975 percentiles for each parameter of interest. 300,000 iterations were generated and the last 100,000 iterations were used to estimate the posterior distributions. For each simulation, 3 chains were iterated to facilitate convergence diagnostics and thinned by 3. Convergence was confirmed visually (density, trace and autocorrelation plots) as well as by calculating the Gelman-Rubin diagnostic (all 1).

The sensitivity of the prevalence estimation was assessed by using alternative prior distributions for the prevalence ratio of calves and yearlings to adults (Appendix S5) and by using a vague prior for the 2017 prevalence. We tested the impact of the dSe being based on short (2 years) versus long (3 years) disease progression. If the mortality of individuals is taken into account, as well as a potential increase in recently infected individuals, the dSe is likely to be better described by an exponential decay of infection times, rather than assuming a random time of infection with equal probability along the whole timeline. Hence, we also tested an assumption with infection times exponentially decaying (excluding any times > 4 yr), and assumed that the mean time since infection was one year. The models using 3 age classes were compared with a model including adults only. Prevalence estimates were also compared with a similar model using a binomial instead of a hypergeometric distribution.

Application 2: Estimating surveillance sensitivity

The scenario-tree model (Fig. 2) was utilized to estimate the surveillance sensitivity for the period 2016 – 2017 for two adjacent herds of Nordfjella zone 1.

We ignored all calves tested (because of the long time period before PrP^{CWD} is found in the lymph nodes and the resulting low dSe of calves) and assumed that adults had three times the risk of being infected compared to yearlings. The relative risk was adjusted for the population proportion of the age classes (Martin, Cameron & Greiner 2007). In the simulations (1000 iterations), we randomly drew the time since infection for each of the tested individuals, and the resulting diagnostic sensitivity for individual *i* in the simulation *j* (dSe_{ij}) was determined by the pathway of individual *i* through the scenario tree and weighted by the adjusted relative risk of the age class. The pathway of an individual through the tree was randomly drawn according to probabilities specified by the testing regime.

We used the hypergeometric distribution, assuming a finite and relatively small population compared to the number of samples tested. Let $ProbAllSampleNeg_j$ denote the probability that there were no test-positive animals found in iteration *j* of *n* samples for a given design prevalence, *p*. According to MacDiarmid (1987), a hypergeometric distribution can be approximated by:

$ProbAllSampleNeg_j = (1 - \sum_{i=1}^{n} dSe_{ij} / PopSize)^{p*PopSize}$

For each of 1000 iterations, we then calculated the surveillance sensitivity as 1 - $ProbAllSampleNeg_j$. The design prevalence was defined by the number of infected animals (threshold to be detected) and the population size. The final distribution of estimated surveillance sensitivities was then summarized as the mean with 95 % confidence intervals. The procedure was repeated for a scenario with a 3-year infection dynamic and for a scenario with an exponential decay distribution of time since infection.

Results

Diagnostic sensitivity

The probability of detecting an infected individual as positive, dSe, was strongly dependent on the time since infection, age class and tissue type, as expected, due to the long duration of infection. When testing brainstem only, the average dSe (including only the detectable stages) was 82 %, 73 % and 0 % for adults, yearlings and calves, respectively, for the parameters and uncertainties chosen (Table 1, Fig. 3). The dSe increased when RLN tissue was included, giving an average dSe of 88 %, 85 % and 45 % for adults, yearlings and calves, respectively (Fig. 3D). There was substantial variation in the distribution of dSe, arising mainly from the different detectability along the timeline of infection. The distribution of dSe was skewed towards low sensitivities, potentially below 0.6 even in adults and yearlings. Additionally, among the simulated animals known to be infected, the probability of infected individuals being in non-detectable stages (i.e. stage 0 and 1 if only brain tissue was tested and stage 0 when RLN was included) decreased from 100 % to 66 % in calves, 52 % to 20 % in yearlings, and from 36 % to 14 % in adults when both RLN and brain tissue were included in the sample (Fig. 3C).

When assuming a 3-year compared to the above 2-year disease development, dSe was approximately unchanged for adults, while it decreased for yearlings and calves since a higher proportion would be in early stages and hence would not be detectable (Appendix S6). When using exponential decay of infection times, dSe decreased for both sample types and all age classes due to the increased proportion of individuals in early stages (Appendix S6), while the probability of infected adults being in stage 1 increased from 23 % to 28 %. As a result, the variation of dSe also increased, with a higher probability of dSe below 0.6.

Application 1: Prevalence in the CWD-infected reindeer herd in Nordfjella

Our model estimated both true (TP) and apparent (AP) prevalence for calves, yearlings and adults and separated in the infection stages 0 to 3 (Appendix S7). The TP was 1.6 % (0.7 %, 3.1 %) in adult reindeer, while the AP was 1.2 % (0.5 %, 2.1 %) in Nordfjella zone 1 in 2017 (Fig. 4). The prevalence estimation was robust to changes in the incubation period, while assuming decay in the distribution of infection times increased the TP only, due to more individuals in early stages. The variation in the prior-distribution for the demographic pattern of infection also had limited impact on the results (results not shown). However, as no calves or yearlings tested positive in 2016 and 2017, there was increased variance in the prevalence estimates.

Application 2: Surveillance sensitivity of adjacent reindeer herds

In 2017, only 9 % of the samples from Hardangervidda and as much as 92 % of the samples from Nordfjella zone 2 included RLN (Appendix S4). Due to the different population sizes, a design prevalence of 5 and 20 infected individuals would represent 0.05 % and 0.2 % prevalence in Hardangervidda and 1 % and 4 % prevalence in Nordfjella zone 2. Combining 2016 and 2017 data, the likelihood of having detected CWD in the larger Hardangervidda was 34 % (95 % CI: 33-35 %) and 57 % (52-61 %) in the smaller Nordfjella zone 2, if assuming five infected individuals (Fig. 5, Appendix S8). If assuming 20 CWD-infected individuals, this increased to 78 % (77-80 %) and 96 % (95-98 %) for Hardangervidda and Nordfjella zone 2, respectively. By extending the disease progression to 3 years, the surveillance sensitivity is unchanged or slightly reduced for both populations. Assuming decay in the distribution of time since infection (increasing proportion in early stages) lowered the surveillance sensitivity (Fig. 5).

Discussion

Surveillance is a crucial first step to combat emerging diseases (Holmes, Rambaut & Andersen 2018). Diseases with long latent stages, slow disease development and a prolonged period of low prevalence are difficult to detect and appropriately estimate the prevalence in early epidemic stages. Differential detectability during disease progression utterly complicates the estimation processes. Here, we approach such challenges posed by the recent emergence of CWD in wild reindeer in Norway, regarded among the top 15 most important emerging issues in biodiversity research globally (Sutherland et al., 2018).

Accounting for differential detection during infection

Imperfect detection remains a limiting factor for assessments of infection levels and disease status that are critical for wildlife surveillance, and diagnostic sensitivity (dSe) is the term used for the probability that an infected animal will be detected by testing. dSe is usually estimated by comparing a test to a reference test and using latent class modelling in the absence of a perfect reference test (van Smeden et al., 2013). Such approaches do not account for the dependence of detectability on disease progression, and the non-detectable, early stages are typically ignored. Individual heterogeneity in detectability frequently arises from increasing intensity during the course of infection. Earlier studies have approached this using extensions of occupancy modelling frameworks to account for differential detectability with an increasing pathogen load during infection, with case examples from amphibians (Miller et al., 2012) and birds (Lachish et al., 2012). This involves repeated measurements of individuals and an assumption of the level of individual infection intensity not changing between measurements. In our case, we cannot fit occupancy models due to a lack of repeated measurements. We provide a novel scenario-tree framework accounting for individual heterogeneity in test sensitivity arising from temporal variation in detectability in multiple tissues during a reasonably predictable infection dynamic. Stochastic simulations were utilized to describe

how the times since infection occurred, and thereby dSe, were distributed among infected individuals. We suggest that our approach may be useful for a broader range of diseases with a slow development. For example, several tick-borne infections may initially be identified in tissue from ears or skin for months post-infection, before later systemic stages may be detected in internal organs, while serology may detect antibodies remaining for years post-infection.

We used this scenario-tree model in two related ways: The estimation of age- and infection-stage-specific true prevalence of CWD in a given population with mixed samples (Application 1), and the estimation of the probability of detecting CWD if the disease was present at a low prevalence in adjacent populations (Application 2). Our two applications use the same model of diagnostic sensitivity, but they differ in that the prevalence model is a Bayesian approach utilizing the age-class- and tissue-specific distribution of dSe to go from apparent to true prevalence, while application 2 is a direct extension of the scenario-tree model, incorporating relative risk between age classes and applied to a specific testing regime and design prevalence. A weighted sum of the dSe of individuals tested is utilized in the estimation of the surveillance sensitivity. In our scenario-tree model, individual dSe is estimated as a continuous function of time since infection, although the scenario-tree (Figure 2) is simplified by categorizing infected individuals into four infection stages. Incorporation of individual sensitivity in freedom from disease modelling, for which surveillance sensitivity is a basic component, is an important application of our framework. In this approach, the next step is to combine the surveillance sensitivity of several years to update our probability of populations being free from CWD.

A deviation from the expected proportion of RLN to brainstem positives will indicate that the underlying model of infection dynamics is inaccurate. We reviewed the North American literature, and the predicted 26 % from our model is within the 19-29 % RLN-only positives reported in studies

with large sample sizes (Appendix S9). Note, however, that the predicted proportion of infected in stage 1 may not be directly comparable to empirical observations, for which some individuals were likely to be misclassified due to imperfect dSe. In any case, our current model was broadly accurate based on existing evidence.

Applications, limitations and further possibilities

Modelling always requires choices regarding simplicity versus necessity of including biological details, and here, we assess the specifics of the case of CWD.

We lack data on demographic patterns of CWD infection in reindeer in Norway and therefore set broad age-specific priors for calves, yearlings and adults based on American studies of CWD in deer (Potapov et al., 2013; Samuel & Storm 2016). Our estimation was robust to changes in prior distributions used for details in demographic infection patterns (results not shown). Adult males often have a 2-3 times higher infection rate in deer, and adding sex-specificity and further age classes would be a natural and technically simple extension of the current application in a North American setting with larger sample sizes (Potapov et al., 2013; Samuel & Storm 2016).

The levels of susceptibility towards CWD and the duration of disease development depend on polymorphisms in the prion protein, in turn determined by mutations in the gene termed *PRNP* (Robinson et al., 2012). Our model readily accounts for variation in the incubation period, as we demonstrate by varying it from 2 to 3 years or with a mix of these. Reindeer have, in general, a larger variation in *PRNP* (8 types) compared to other cervids (Robinson et al., 2012). The two transmission studies inoculating CWD in captive reindeer with different *PRNP* genotypes suggested that 2-3 years were likely to capture the main variation in disease duration (Mitchell et al., 2012; Moore et al., 2016). This is in agreement with one CWD-positive 2.5-year-old reindeer from Nordfjella that had a very high obex score, indicating a late stage of disease. With more empirical data, the model can be set to incorporate known proportions of different *PRNP* genotypes in a population affecting the

susceptibility and incubation period, but the result of our modelling suggested that the duration of infection did not affect the estimated prevalence at the current low prevalence levels notably (Fig. 4).

Another issue for estimation is that there may be a higher proportion of individuals in early stages of infection if there is an epidemic growth of disease and/or (increased) mortality over the time of infection. We simulated such effects by adding an exponential decay of time since infection instead of drawing individuals randomly along the infection timeline, as in our main model. This had a notable effect on the estimation of true prevalence, by increasing the variance (Fig. 4). The higher proportion of individuals in early, non-detectable stages further markedly lowers the surveillance sensitivity (Fig. 5). With more data, possibly the force of infection approaches (Heisey et al., 2010) or SEIR-modelling can explicitly account for temporal dynamics of infected individuals depending on R0 and mortality, and possibly also account for a potential seasonal pattern of infection.

Tradition of citizen science involving hunters

From 2018, a surveillance program for CWD is being implemented by the European Union in all countries having populations of moose and/or reindeer, including Sweden, Finland, Estonia, Lithuania, Latvia and Poland (EFSA Panel on Biological Hazards (BIOHAZ) et al., 2016). All surveillance programs have to incorporate decisions regarding the priority for determining how many areas to cover (Russell et al., 2015). The citizen science approach is often used as a strategy to obtain wildlife data from large geographic areas, but it is important to know how it may affect sample quality and cause bias (Steger, Butt & Hooten 2017). In Scandinavia, hunters are part of a well-developed system for population monitoring, and hunters are used to gather tissue samples for scientific purposes. However, the willingness to participate depends on a number of factors. In Nordfjella, most hunting areas have proximity to roads, and it was feasible to bring the whole head to a sampling station where both obex and RLN could be harvested by a professional. In contrast, large parts of Hardangervidda are more remote and often require many hours of hiking to reach a road;

additionally, this area had no organized sampling stations. The decision regarding whether to ask the hunter to carry the head (2-7 kg), which would result in few samples, or to allow hunters to take more samples at the site, despite the probable loss of sample quality, required consideration. For most untrained personnel, RLN represent a more challenging tissue to localize, and subsequent collection can be less successful, but the absence of RLN in the sample will decrease the test sensitivity, especially in the first stage, when PrP^{CWD} is only detectable in RLN. Our approach allows a unified estimation even in cases with different tissue types arising from a citizen science-based sampling approach.

Acknowledgements

We are grateful to many sources that contributed to this work, including Kåre Rudningen, Tor Bakken, thousands of hunters for gathering samples, the Norwegian Environment Agency and the Ministry of Agriculture and Food for funding, and the pathology department of the Norwegian Veterinary Institute for the excellence of their work with prion testing. We are grateful for comments to a previous draft by two anonymous referees and the editor.

Authors' contributions

AM & HV initiated the project. HV developed the main conceptual idea and the method with input from PH, SB, ST, JV, CMR and AM. EBM and HV implemented the reindeer population model. JV, CMR and PH organized the CWD data collection and database management. SB is responsible for testing for CWD. OS organized the reindeer surveillance data collection. AM and HV drafted the manuscript with the methods parts contributed by PH, CMR and EBM. All authors read and improved subsequent versions.

Data Accessibility

The reindeer data and scripts used in this manuscript are available from the Dryad Digital Repository with DOI: http// doi:10.5061/dryad.q84p862 (Viljugrein et al., 2018).

References

Benestad, S.L., Mitchell, G., Simmons, M., Ytrehus, B., & Vikøren, T. (2016) First case of chronic wasting disease in Europe in a Norwegian free-ranging reindeer. *Veterinary Research*, **47,** 88.

Davenport, K.A., Hoover, C. E., Bian, J., Telling, G. C., Mathiason, C. K., & Hoover, E. A. (2017) PrPC expression and prion seeding activity in the alimentary tract and lymphoid tissue of deer. *Plos One*, **12,** e0183927.

Davenport, K.A., Mosher, B. A., Brost, B. M., Henderson, D. M., Denkers, N. D., Nalls, A. V., McNulty, E., Mathiason, C. K., & Hoover, E. A. (2018) Assessment of Chronic Wasting Disease prion sheeding in deer saliva with occupancy modeling. *Journal of Clinical Microbiology*, **56,** e01243-17.

Delignette-Muller, M.L. & Dutang, C. (2015) fitdistplus: An R package for fitting distributions. *Journal of Statistical Software*, **64,** 1-34.

Doherr, M.G. & Audigé, L. (2001) Monitoring and surveillance for rare health-related events: a review from the veterinary perspective. *Philosophical Transactions of the Royal Society Series B*, **356,** 1097-1106.

EFSA Panel on Biological Hazards (BIOHAZ), Ricci, A., Allende, A., Bolton, D., Chemaly, M., Davies, R., Escámez, P. S. F., Gironés, R., Herman, L., Koutsoumanis, K., Lindqvist, R., Nørrung, B., Robertson, L., Sanaa, M., Skandamis, P., Snary, E., Speybroeck, N., Kuile, B. T., Threlfall, J., Wahlström, H., Benestad, S., Gavien-Widen, D., Miller, M. W., Ru, G., Telling, G. C., Tryland, M., Pelaez, A. O., & Simmons, M. (2016) Chronic Wasting Disease (CWD) in cervids. *EFSA Journal*, **15,** 4667.

European Food Safety Authority (EFSA) (2005) Scientific report of the European Food Safety Authority on the evaluation of rapid post mortem TSE Tests intended for small ruminants. *EFSA Journal*, **3,** 49r-n/a.

Fox, K.A., Jewell, J. E., Williams, E. S., & Miller, M. W. (2006) Patterns of PRPCWD accumulation during the course of Chronic Wasting Disease infection in orally inoculated mule deer (*Odocoileus hemionus*). *Journal of General Virology*, **87,** 3451-3461.

Funk, S., Bogich, T. L., Jones, K. E., Kilpatrick, A. M., & Daszak, P. (2013) Quantifying trends in disease impact to produce a consistent and reproducible definition of an emerging infectious disease. *Plos One*, **8,** e69951.

Haley, N.J. & Richt, A. J. (2017) Evolution of diagnostic tests for Chronic Wasting Disease, a naturally occurring prion disease of cervids. *Pathogens*, **6,** E35.

Heisey, D.M., Osnas, E. E., Cross, P. C., Joly, D. O., Langenberg, J. A., & Miller, M. W. (2010) Linking process to pattern: estimating spatiotemporal dynamics of a wildlife epidemic from cross-sectional data. *Ecological Monographs*, **80,** 221-240.

Hibler, C.P., Wilson, K. L., Spraker, T. R., Miller, M. W., Zink, R. R., DeBuse, L. L., Andersen, E., Schweitzer, D., Kennedy, J. A., Baeten, L. A., Smeltzer, J. F., Salman, M. D., & Powers, B. E. (2003) Field validation and assessment of an enzyme-linked immunosorbent assay for detecting chronic wasting disease in mule deer (*Odocoileus hemionus*), white-tailed deer (*Odocoileus virginianus*), and Rocky Mountain elk (*Cervus elaphus nelsoni*). *Journal of Veterinary Diagnostic Investigation*, **15,** 311-319.

Holmes, E.C., Rambaut, A., & Andersen, K. G. (2018) Pandemics: spend on surveillance, not prediction. *Nature*, **558,** 182.

Hoover, C.E., Davenport, K. A., Henderson, D. M., Denkers, N. D., Mathiason, C. K., Soto, C., Zabel, M. D., & Hoover, E. A. (2017) Pathways of prion spread during early Chronic Wasting Disease in deer. *Journal of Virology*, **91,** e00077-17.

Jones, K.E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008)

Global trends in emerging infectious diseases. *Nature*, **451,** 990-993.

Keane, D.P., Barr, D. J., Keller, J. E., Hall, S. M., Langenberg, J. A., & Bochsler, P. N. (2008) Comparison of retropharyngeal lymph node and obex region of the brainstem in detection of Chronic Wasting Disease in white-tailed deer (*Odocoileus virginianus*). *Journal of Veterinary Diagnostic Investigation*, **20,** 58-60.

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.

Lewis, F.I. & Torgerson, P. R. (2012) A tutorial in estimating the prevalence of disease in humans and animals in the absence of a gold standard diagnostic. *Emerging Themes in Epidemiology*, **9,** article 9.

Lunn, D., Jackson, C., Best, N., Thomas, A., & Spegelhalter, D. (2013) *The BUGS book: A practical introduction to Bayesian analysis.* CRC Press, Boca Raton, Florida.

MacDiarmid, S.C. (1987) A theoretical basis for the use of a skin test for brucellosis surveillance in extensively-managed cattle herds. *Revue Scientific et technique - Office Internatinal des epizooties*, **6,** 1029-1035.

Martin, P.A., Cameron, A. R., & Greiner, M. (2007) Demonstrating freedom from disease using multiple complex data sources. *Preventive Veterinary Medicine*, **79,** 71-97.

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.

Miller, M.W., Williams, E. S., McCarty, C. W., Spraker, T. R., Kreeger, T. J., Larsen, C. T., & Thorne, E. T. (2000) Epizootiology of Chronic Wasting Disease in free-ranging cervids in Colorado and Wyoming. *Journal of Wildlife Diseases*, **36,** 676-690.

Mitchell, G.B., Sigurdson, C. J., O'Rourke, K. I., Algire, J., Harrington, N. P., Walther, I., Spraker, T. R., & Balachandran, A. (2012) Experimental oral transmission of Chronic Wasting Disease to

reindeer (*Rangifer tarandus tarandus*). *Plos One*, **7,** e39055.

Moore, S.J., Kunkle, R., Greenlee, M. H. W., Nicholson, E., Richt, J., Hamir, A., Waters, W. R., & Greenlee, J. (2016) Horizontal transmission of Chronic Wasting Disease in reindeer. *Emerging infectious diseases*, **22,** 2142.

Potapov, A., Merrill, E., Pybus, M., Coltman, D., & Lewis, M. A. (2013) Chronic wasting disease: possible transmission mechanisms in deer. *Ecol Mod*, **250,** 244-257.

Rhyan, J.C. & Spraker, T. R. (2010) Emergence of disease from wildlife reservoirs. *Veterinary Pathology*, **47,** 34-39.

Robinson, S.J., Samuel, M. D., O'Rourke, K. I., & Johnson, C. J. (2012) The role of genetics in chronic wasting disease of North American cervids. *Prion*, **6,** 153-162.

Russell, R.E., Gude, J. A., Anderson, N. J., & Ramsey, J. M. (2015) Identifying priority chronic wasting disease surveillance areas for mule deer in Montana. *Journal of Wildlife Management*, **79,** 989-997.

Samuel, M.D. & Storm, D. J. (2016) Chronic wasting disease in white-tailed deer: infection, mortality, and implications for heterogeneous transmission. *Ecology*, **97,** 3195-3205.

Steger, C., Butt, B., & Hooten, M. B. (2017) Safari science: assessing the reliability of citizen science data for wildlife surveys. *Journal of Applied Ecology*, **54,** 2053-2062.

Sutherland, W.J., Butchart, S. H. M., Connor, B., Culshaw, C., Dicks, L. V., Dinsdale, J., Doran, H., Entwistle, A. C., Fleishman, E., Gibbons, D. W., Jiang, Z., Keim, B., Roux, X. L., Lickorish, F. A., Markillie, P., Monk, K. A., Mortimer, D., Pearce-Higgins, J. W., Peck, L. S., Pretty, J., Seymour, C. L., Spalding, M. D., Tonneijck, F. H., & Gleave, R. A. (2018) A 2018 horizon scan of emerging Issues for global conservation and biological diversity. *Trends in Ecology & Evolution*, **33,** 47-58.

Thomsen, B.V., Schneider, D. A., O'Rourke, K. I., Gidlewski, T., McLane, J., Allen, R. W., McIsaac, A. A., Mitchell, G. B., Keane, D. P., Spraker, T. R., & Balachandran, A. (2012) *Diagnostic accuracy of rectual 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. USDA National Wildlife Research Center - Staff Publications 1200, University of Nebraska, Lincoln.

van Smeden, M., Naaktgeboren, C. A., Reitsma, J. B., Moons, K. G. M., & de Groot, J. A. H. (2013) Latent class models in diagnostic studies when there is no reference standard - a systematic review. *American Journal of Epidemiology*, **179,** 423-431.

Viljugrein, H., Hopp, P., Benestad, S. L., Nilsen, E. B., Våge, J., Tavornpanich, S., Rolandsen, C. M., Strand, O., & Mysterud, A. (2018) Data from: A method that accounts for differential detectability in mixed samples of long-term infections with applications to the case of Chronic Wasting Disease in cervids. *Methods in Ecology and Evolution*, doi:10.5061/dryad.q84p862

Williams, E.S. & Young, S. (1980) Chronic wasting disease of captive mule deer: a spongiform encephalopathy. *Journal of Wildlife Diseases*, **16,** 89-98.

Supporting Information

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

Appendix S1. Study areas, census data and population model for reindeer.

Appendix S2. Details on the test regime.

Appendix S3. A table with test sensitivity (dSe).

Appendix S4. A table with the sample size of CWD tested individuals and the proportion tested for lymph nodes from the total number.

Appendix S5. The prior distributions of the prevalence ratios between age classes.

Appendix S6. A figure of the distribution of the diagnostic sensitivity (dSe) for individuals from each of the three age classes.

Appendix S7. A table with the estimated prevalence of CWD in the reindeer population in Nordfjella zone 1, Norway.

Appendix S8. A table with the likelihood of having detected CWD in the reindeer populations in

Nordfjella zone 2 and Hardangervidda, Norway.

Appendix S9. An overview of studies testing for CWD in retropharyngeal lymph nodes (RLN) and obex.

Table 1. The parameters used for test sensitivity are based on a combination of published information and expert judgements. Note that the lower analytical test sensitivity of RLN and obex combined is due to a lower sample mass. Diagnostic sensitivity of a tissue sample is dependent on the tissue type (RLN, obex of low or high quality) and is a continuous function of disease progression according to Figure 1. Diagnostic sensitivity is modelled by scaling scores above 1 in Figure 1 between the lower and upper levels as given by SeLow and the analytical test sensitivity, respectively. Sensitivity for scores less than 1 is set to zero. Uncertainty around the two parameters, PrLQ and Se_tissue, is modelled as a beta-pert distribution given by the expected value, min and max.

Term of interest	Parameter	Expected (min-max)	References
Analytical test sensitivity obex	Se_obex	0.925	(Hibler et al., 2003)
Analytical test sensitivity RLN	Se_RLN	0.988	(Hibler et al., 2003)
Analytical test sensitivity obex+RLN	Se_RLN & obex *	0.95	expert opinion
Proportion of low-quality obex samples	PrLQ	0.22 (0.02-0.6)	expert opinion
Parameters defining lower level of mean Se_tissue	SeLow_RLN	0.25	expert opinion
	SeLow_ObexLQ	0.25	expert opinion
	SeLow_ObexHQ	0.5	expert opinion
Tissue/sample sensitivity	Se_tissue_i	Se_tissue_i (Se_tissue_i^var_low-Se_tissue_i^var_up)	function of tissue type and time since infection *i*
Parameters defining variation around mean Se_tissue	var_low	1.25	expert opinion
	var_up	0.6	expert opinion

***Note: some pooling of samples in first testing, cfr. Appendix.**

Figure captions

Figure 1. Association between the assumed disease progression of Chronic Wasting Disease and the detectability of prions in lymph nodes (RLN, blue line), and brain tissue (obex and brainstem, black line), starting at the time of exposure to infection. The duration of the development varied from 24 to 36 months in the modelling, but the shape of the curves was constant.

Figure 2. Stochastic scenario tree representing the testing regime for Chronic Wasting Disease in a reindeer population/herd. The pathway of simulated individuals through the scenario tree is determined by time since infection (randomly drawn), population proportions and probabilities regarding the testing regime. The probability of detecting CWD (diagnostic sensitivity) for a given age class, stage of infection, sample of lymph nodes (RLN) and obex tissue of various quality is modelled as a nonlinear function of time since infection (according to Figure 1 and assumptions as specified in Table 1).

Figure 3. Estimated diagnostic sensitivities (dSe), shown for the assumption of **A**) 2-yr and **B**) 3-yr infection periods, vary according to time since infection and depend on the tissue (lymph nodes, RLN and/or brain tissue, obex). A higher dSe indicates later stages of the CWD infection period. **C)** The proportion of non-detectable (dSe=0) CWD-infected calves, yearlings and adults depending on the sample type. **D**) The dSe of calves, yearlings and adults when including both RLN and brain tissue compared to only brain tissue. Sensitivity distributions (dSe>0) are summarized by the mean, median and the 95 % confidence interval.

Figure 4. The estimated apparent (AP) and true (TP) prevalence in the CWD-infected reindeer herd in Nordfjella zone 1, Norway, in 2017 (the mean and 95 % credible intervals). AP and TP are given for the overall population (weighted mean), for adults, and for adults in different infection stages under 3 different assumptions (see text for details).

Figure 5. Surveillance sensitivity for design prevalences assuming 5, 10, 20 or 50 infected individuals in a small ("N" - Nordfjella zone 2 with ~500 reindeer) and a large population ("H" - Hardangervidda with ~10,000 individuals), Norway.