1-2008

Sampling Considerations for Disease Surveillance in Wildlife Populations

Sarah M. Nusser
Iowa State University, nusser@iastate.edu

William R. Clark
Iowa State University

David L. Otis
United States Geological Survey

Ling Huang
Iowa State University, lingh@iastate.edu

Sampling Considerations for Disease Surveillance in Wildlife Populations

Abstract
Disease surveillance in wildlife populations involves detecting the presence of a disease, characterizing its prevalence and spread, and subsequent monitoring. A probability sample of animals selected from the population and corresponding estimators of disease prevalence and detection provide estimates with quantifiable statistical properties, but this approach is rarely used. Although wildlife scientists often assume probability sampling and random disease distributions to calculate sample sizes, convenience samples (i.e., samples of readily available animals) are typically used, and disease distributions are rarely random. We demonstrate how landscape-based simulation can be used to explore properties of estimators from convenience samples in relation to probability samples. We used simulation methods to model what is known about the habitat preferences of the wildlife population, the disease distribution, and the potential biases of the convenience-sample approach. Using chronic wasting disease in free-ranging deer (*Odocoileus virginianus*) as a simple illustration, we show that using probability sample designs with appropriate estimators provides unbiased surveillance parameter estimates but that the selection bias and coverage errors associated with convenience samples can lead to biased and misleading results. We also suggest practical alternatives to convenience samples that mix probability and convenience sampling. For example, a sample of land areas can be selected using a probability design that oversamples areas with larger animal populations, followed by harvesting of individual animals within sampled areas using a convenience sampling method.

Keywords
Ecology Evolution and Organismal Biology, chronic wasting disease, disease detection, disease prevalance, sample design, surveillance, waiting time distribution

Disciplines
Biometry | Biostatistics | Population Biology

Comments
This article is from *Journal of Wildlife Management* 72 (2008): 52–60, doi:10.2193/2007-317.

Special Section: Observational Studies

Sampling Considerations for Disease Surveillance in Wildlife Populations

SARAH M. NUSSER,[1] *Department of Statistics and Center for Survey Statistics and Methodology, 222 Snedecor Hall, Iowa State University, Ames, IA 50011-1210, USA*

WILLIAM R. CLARK, *Department of Ecology, Evolution, and Organismal Biology, 233 Bessey Hall, Iowa State University, Ames, IA 50011-1020, USA*

DAVID L. OTIS, *United States Geological Survey, Iowa Cooperative Fish and Wildlife Research Unit, 342 Science II, Iowa State University, Ames, IA 50011-3221, USA*

LING HUANG, *Department of Statistics, 104 Snedecor Hall, Iowa State University, Ames, IA 50011-1210, USA*

ABSTRACT Disease surveillance in wildlife populations involves detecting the presence of a disease, characterizing its prevalence and spread, and subsequent monitoring. A probability sample of animals selected from the population and corresponding estimators of disease prevalence and detection provide estimates with quantifiable statistical properties, but this approach is rarely used. Although wildlife scientists often assume probability sampling and random disease distributions to calculate sample sizes, convenience samples (i.e., samples of readily available animals) are typically used, and disease distributions are rarely random. We demonstrate how landscape-based simulation can be used to explore properties of estimators from convenience samples in relation to probability samples. We used simulation methods to model what is known about the habitat preferences of the wildlife population, the disease distribution, and the potential biases of the convenience-sample approach. Using chronic wasting disease in free-ranging deer (*Odocoileus virginianus*) as a simple illustration, we show that using probability sample designs with appropriate estimators provides unbiased surveillance parameter estimates but that the selection bias and coverage errors associated with convenience samples can lead to biased and misleading results. We also suggest practical alternatives to convenience samples that mix probability and convenience sampling. For example, a sample of land areas can be selected using a probability design that oversamples areas with larger animal populations, followed by harvesting of individual animals within sampled areas using a convenience sampling method. (JOURNAL OF WILDLIFE MANAGEMENT 72(1):52–60; 2008)

DOI: 10.2193/2007-317

KEY WORDS chronic wasting disease, disease detection, disease prevalence, sample design, surveillance, waiting time distribution.

Surveillance of zoonotic diseases has become increasingly important to wildlife ecologists and managers. Disease surveillance in wildlife populations involves detecting the presence of the disease, estimating its prevalence and spatial distribution, and monitoring its progression. Probability sampling methods appropriate for achieving surveillance objectives are unfamiliar to many wildlife specialists, in part because training for most wildlife biologists is based on classical agricultural experimental design for which control, manipulation, and replication are cornerstones (Otis 2001, Shaffer and Johnson 2008). In addition, biologists most often learn about estimation only under the most simple survey designs such as simple random sampling (SRS) and stratified random sampling (STS), and they often do not learn when to apply appropriate estimators (Taylor et al. 2000).

Sample surveys rely on probability sampling to choose sample units (e.g., small areas or individuals) for observation from a population of interest. Probability sampling is most often used when the objective is to estimate population means and totals such as disease prevalence and the number of diseased animals in a population (Lohr 1999, Thompson 2002), and it is commonly used by government agencies to monitor change in natural resource settings (Edwards 1998, Nusser et al. 1998, Olsen et al. 1999). In probability sampling, well-defined rules are invoked to randomly select a sample of units for observation from a list of all possible sample units (i.e., sampling frame) that conceptually represent the population of interest. When sampling wildlife populations, often the sampling frame is a list of mutually exclusive areas (e.g., plots of land) containing individual animals (i.e., population elements). The selection rules and sampling frame allow us to calculate the probability that a sampling unit and (or) population element is included in the sample and to derive estimators appropriate for a specific sample design. Because the selection probabilities are known, valid statistical properties of estimators can be derived, which provide the basis for evaluating the scientific credibility of the estimates.

Although most wildlife ecologists are familiar with concepts of SRS, observations are often obtained by various ways that are convenient, but certainly not random (e.g., road kills, hunter-shot samples). When convenience sampling is used, selection probabilities cannot be described analytically, which makes it impossible to derive statistically valid estimators or appropriate standard errors. In practice, untested and typically unfounded assumptions are made about how well the convenience sample reflects the population (e.g., the convenience sample approximates a SRS sample).

The focus of this special section is on how to deal with the practical need to rely on observational data in studies of wildlife populations. Other papers focus on issues such as randomization of treatments to assess possible cause-and-effect relationships within a study population (Shaffer and Johnson 2008). We are interested in describing the

[1] *E-mail: nusser@iastate.edu*

characteristics of a population at a point in time or over a period of time (e.g., estimating disease prevalence). This kind of problem requires that data be collected from a statistically representative sample (i.e., a probability sample that can be used to make valid estimates and inferences about the population). Just as experimental design is the rigorous method of studying the effect of a treatment on a population, survey sampling is a highly reliable and well-established method for estimating population characteristics. Besides design-based estimation, model-based estimators (e.g., incorporating geospatial correlations and trends) or those based on combinations of model- and design-based estimators are also effective approaches (Ver Hoef 2002).

Although the design of wildlife population sampling has substantially improved in the last 25 years, sampling designs for disease surveillance and other field investigations often have used nonprobability (e.g., convenience) selection. Wildlife researchers recognize that SRS is often impractical, and they instinctively use ad hoc devices that mimic more complex sample designs. For example, researchers may sample from separate segments of the population, which is like stratifying a probability sample, or may spend relatively more effort on collecting observations from informative units (e.g., animals exhibiting disease symptoms), which is similar to the concept of unequal probability sampling. In doing so, wildlife specialists use prior knowledge (e.g., that the animal population exhibits habitat preference) to guide the observational process towards a seemingly more representative population estimate. Less often do researchers consider optimally allocated sample sizes or compare efficiency of alternative designs given available resources (Lohr 1999, Thompson 2002).

We discuss the opportunities and difficulties of using rigorous sampling methods for wildlife disease surveillance, which has received great attention recently, as exemplified by concerns about potential impact of West Nile Virus (WNV; Eidson et al. 1999). After the initial outbreak of WNV in birds and humans around New York City, a passive sampling approach of testing conveniently collected dead birds was used to broadly characterize the spread of WNV. However, agencies recognize that if they conduct more rigorously designed disease surveillance of wildlife populations, they are more likely to detect and understand the epizootiology of infectious and zoonotic diseases and are therefore better prepared to protect wildlife, domestic animals, and human populations (Mörner et al. 2002). Early detection of disease is essential to prompt response, but making a valid statement about the presence, and especially the absence, of a disease or the expected waiting time until a disease is detected is extremely difficult because of the statistical properties of detecting rare events (Doherr and Audige 2001,Venette et al. 2002).

We use a simple example from the surveillance of chronic wasting disease (CWD) to illustrate issues associated with convenience and probability sampling. Chronic wasting disease is an infectious neurological disease of North American cervids. The disease is similar to scrapie, bovine spongiform encephalopathy, and a human form known as Creutzfeldt-Jakob disease (Williams et al. 2002). The disease has been known from populations of mule deer (*Odocoileus hemionus*) in Colorado and Wyoming, USA, for decades (Miller et al. 2000), but when the disease was detected in white-tailed deer (*O. virginianus*) in Wisconsin, USA, many state and federal agencies were stimulated to plan and conduct surveillance (Samuel et al. 2003). Prevalence of CWD generally was assumed to be low (0.5–1.5%), although sampling has subsequently revealed rates estimated to be as great as 10–11% on a local basis (D. O. Joly, United States Geological Survey, unpublished data). The distribution of the disease and the transmission mechanisms are poorly known (Courchamp et al. 2000). Agencies began surveillance with convenience samples of deer that were easily accessed, especially hunter-shot deer at check stations or meat processing lockers, road-killed deer, and occasionally by sharp shooting. Some surveillance was conducted by collecting from presumed sick animals exhibiting end-stage CWD symptoms such as poor body condition. These sampling approaches assumed sampled deer were typical for the larger population being studied. Most wildlife agencies initially assumed the disease was uniformly or randomly distributed spatially when they planned surveillance sampling (Beringer et al. 2003, Diefenbach et al. 2004). Investigators also made estimates of required sample sizes and associated confidence levels assuming SRS for sampling and random disease distribution (Schmitt et al. 1997, Miller et al. 2000).

We demonstrate the use of simulation to explore how the properties of estimators obtained from alternative probability and convenience sample designs for disease surveillance may vary. Our modeling approach applies relevant species life history characteristics in the context of a specific landscape. By applying alternative sample designs to simulated animal populations, we demonstrate how estimators are affected by the assumption that the disease is randomly distributed throughout the population and the biases of different sample designs.

METHODS

A wildlife specialist who wishes to obtain a sample of animals for surveillance would begin by synthesizing knowledge of the biology of the target population and the epidemiology of the disease. In practice, a pilot study is rarely feasible, but investigators can use simulations to formally represent prior information as a population and to study the properties of estimators obtained by various surveillance designs. In this section we summarize simulation methods we used to generate populations of animals on a landscape derived from a Geographic Information System (GIS) coverage and to assign a disease status to each animal. We discuss alternative sampling approaches to draw samples from the simulated animal populations for comparing surveillance estimators based on animals harvested under the alternative sample designs.

Generating Animal Populations with Disease Status Indicators

Distributing individual animals on the landscape.—To represent animals on the landscape, we used an algorithm that physically distributed individual animals over the landscape using assumptions that reflected habitat preferences of deer (Huang 2005). The landscape was from Fayette County in northeast Iowa, USA. We represented the landscape using the United States Geological Survey 1992 National Land Cover Data, a geospatial coverage developed from satellite imagery and other resources in which 1 of 21 land cover classes was uniquely assigned to each 30-m pixel in the coverage (U.S. Geological Survey 1992). We considered each land cover class a separate habitat type in our simulation. The landscape consisted primarily of agricultural crops (66% of the land area, almost entirely row crops); pasture, hayland, grassland, and other herbaceous cover (16%); and forest (15%, nearly all deciduous; Fig. 1).

For each simulation replicate, we randomly selected locations of individual animals in a deer population within each habitat type. We derived the number of deer in each Fayette County habitat type by multiplying the surface area for each habitat type by the expected deer densities for each habitat type (W. Suchy, Iowa Department of Natural Resources, personal communication). The deer densities represented conditions during the fall hunting season, which was the primary sampling time for collecting deer samples in Iowa. For example, on average, we expected a fall density of 11.6 deer/km^2 within the 276 km^2 of deciduous forest and 1.9 deer/km^2 within the 1,242 km^2 of row crops. Thus, we expected 3,198 deer to be present in deciduous forest and 2,398 deer to be present in the row-cropped areas. Using this approach, we determined the total number of deer in Fayette County to be 7,000. To create a spatial distribution for the 7,000 deer, a list of pixels for each habitat type was created. For each simulation replicate, a computer program used SRS without replacement to select the corresponding number of pixels for each habitat (i.e., expected no. of deer in the habitat) from the pixel list for the habitat (Fig. 1).

This process was repeated for 1,000 simulation replicates. For each replicate, the 7,000 deer were stored in a file with each deer's pixel location, habitat type, and Public Land Survey section and township code. The file formed the basis for assigning disease status for each of the 2 disease models.

Models for assigning disease status to individual animals.—Because the mechanism by which CWD is spread to other individuals is poorly understood (Courchamp et al. 2000), we considered 2 disease models: a baseline random distribution, and a distribution in which most of the diseased animals were clustered with a few additional outlying diseased individuals (Miller and Conner 2005), which we refer to as the hot spot and spark model. In conducting the simulation, we created 2 representations, one for each disease model, for each of the 1,000 simulation replicates of baseline deer populations.

The random model assumed that the disease distribution did not depend on any factor, such as neighboring infected deer or environmental contamination. This is an unrealistic model, but it provides a benchmark for other models and represents an assumption that is often made when calculating sample sizes. We assumed a 1.5% prevalence rate, based on observed rates in a CWD-infected area of Wisconsin (Joly et al. 2003). For each realization of the animal population, a computer program randomly assigned (using SRS) a positive disease status to 105 of the 7,000 deer inhabiting Fayette County. An uninfected disease status was assigned to remaining deer (Fig. 1a).

For our application of the hot spot and spark model, we assumed that there was roughly a 20% prevalence rate within the concentrated nucleus of infection in a small region (Joly et al. 2003) and that an infected animal occasionally migrated to another area of the region. For each realization of the animal population, a computer program randomly assigned a positive disease status to 100 of the approximately 500 deer in a contiguous 6,200-ha forested area. Only lightly traveled roads passed through the hot spot region, which established a worst-case scenario for road-kill convenience samples. We randomly assigned a positive disease status to 5 of the roughly 6,500 deer located in the remainder of the county to represent diseased migrants (Fig. 1b).

Sample Designs and Metrics for Evaluating Animal Disease Status

Overview.—Our primary interest was in comparing estimators from formal probability sample designs with known properties to estimators from convenience sampling approaches whose properties are not known in practice. We considered representations of 2 convenience sampling approaches that have been used by states: road-kill convenience sampling, in which tissue samples from animals recently killed via vehicular collisions are collected for assays; and hunter-based sampling, in which state representatives ask hunters for permission to take samples from animals they have just shot. We compared these convenience sampling approaches to 2 probability sample designs: a benchmark SRS design and a more sophisticated probability sample design that might be used to implement sharp shooting as part of an organized surveillance campaign conducted by state representatives.

Each of the 4 designs was applied to the 1,000 deer populations generated under the randomly distributed disease model and to the 1,000 deer populations generated under the hot spot and spark disease model. All samples had a sample size of 120 deer. Thus, for each simulation replicate, the procedure generated 8 separate samples of 120 deer, one for each of the 8 conditions (2 disease models $\times$ 4 sample designs).

For each of the 8 simulation conditions, we evaluated 3 parameters that wildlife agencies might use to evaluate disease status: prevalence, detection probability, and waiting time. We defined prevalence as the proportion of the population that is infected. Probability of detection is the probability of sampling $\geq$1 infected individual. Waiting

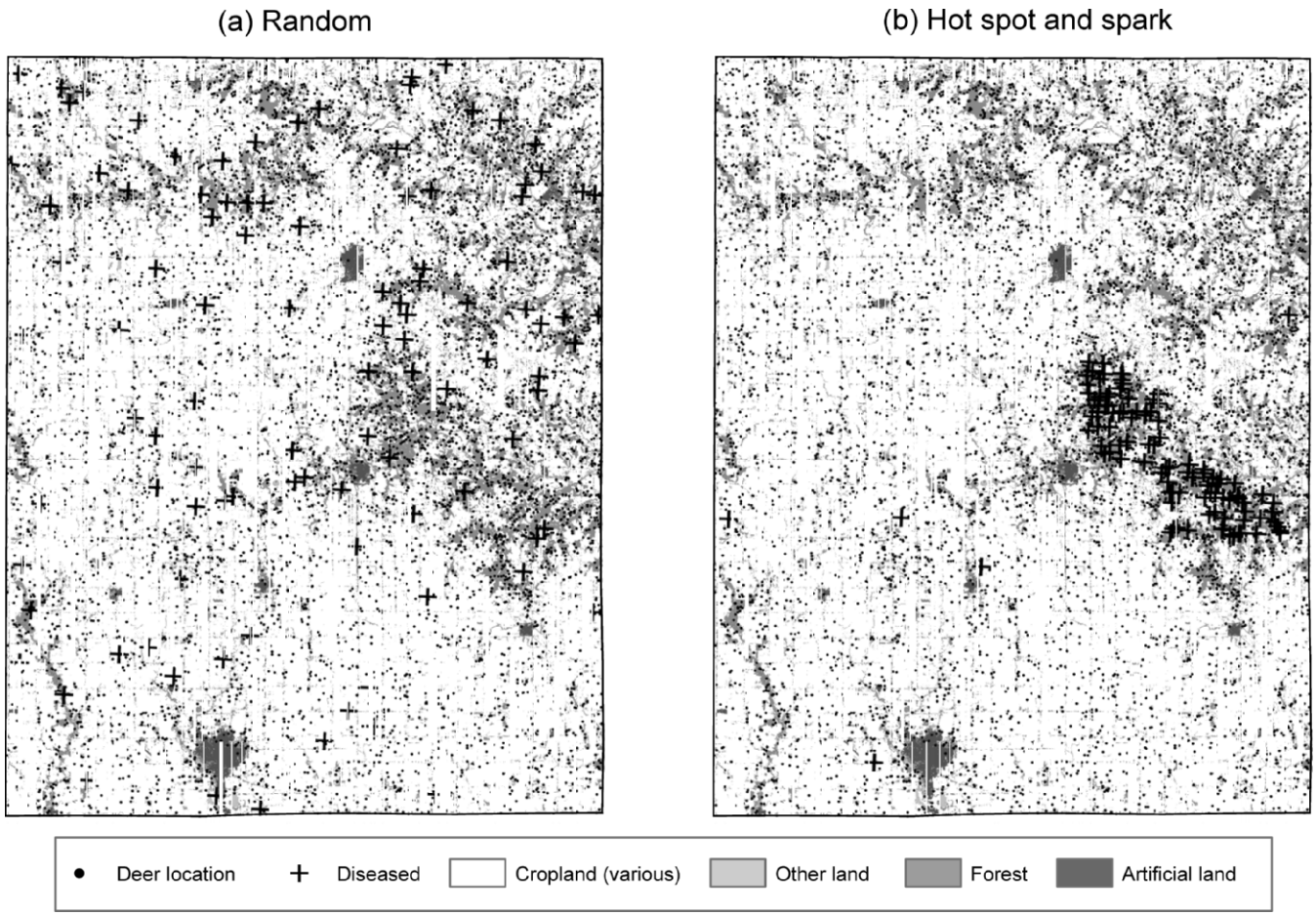

Figure 1. A simulation realization depicting the spatial distribution of healthy and diseased deer in relation to habitat using (a) random and (b) hot spot and spark disease models.

time is the mean number of years until first detection, assuming an annual monitoring program.

Sampling frames.—A sampling frame is a list of units in the population from which a sample is selected. The deer sampling frame for the SRS and 2 convenience sample designs was the list of deer for each generated population, with the pixel coordinate in the coverage, township and section for the pixel, the pixel habitat, and the value of disease status indicator for the deer. For the 2-stage cluster sample for organized sharp shooting, the first-stage frame was a list of sections with township identifier and fraction of forest area for each section. The second-stage frame was a list of deer in each sampled section from the deer frame that was constructed for the other 3 designs.

Probability sample designs.—We studied 2 probability sample designs: SRS and stratified 2-stage cluster sampling. The basic probability design was SRS (Fig. 2a), in which every sample of n animals from a population of N animals had an equal chance of being selected. Simple random sampling represented a benchmark rather than a reasonable alternative design. It is not a practical method of obtaining deer, and we did not expect it to provide gains in statistical efficiency (Lohr 1999).

A more realistic probability sample design for this problem is a stratified 2-stage unequal probability cluster sample, which involves 2 stages of sampling within each stratum. If properly designed, a cluster sampling approach offers a more practical and less costly approach to sampling individual deer by concentrating the harvest of several individuals within a sampled land area and by targeting land areas preferred by the animal population for sampling. Nusser and Klaas (2003) used this approach to assess field accuracy of GIS land cover maps. In our setting, an agency may be interested in sharp-shooting to collect animal samples as part of an organized surveillance effort. We assumed that field staff would go to specific areas to shoot or anesthetize animals to obtain samples. The stratified 2-stage cluster sample design represented an idealized implementation of this approach, where land areas were preselected using a probability design and sharp-shooters gathered samples from 3 randomly selected animals per section. We assumed that sharp-shooters did not exhibit bias in harvesting individual deer within the section, particularly with respect to disease status. In our simulation, we divided the county into 10 geographic strata, defined by 2 contiguous townships (i.e., 6×12 Public Land Survey System sections). In our simulation, the first-stage sampling unit, called a cluster or primary sampling unit, was a section and represented a cluster of individual animals. Within each stratum, there were 72 sections, from which we randomly selected 4

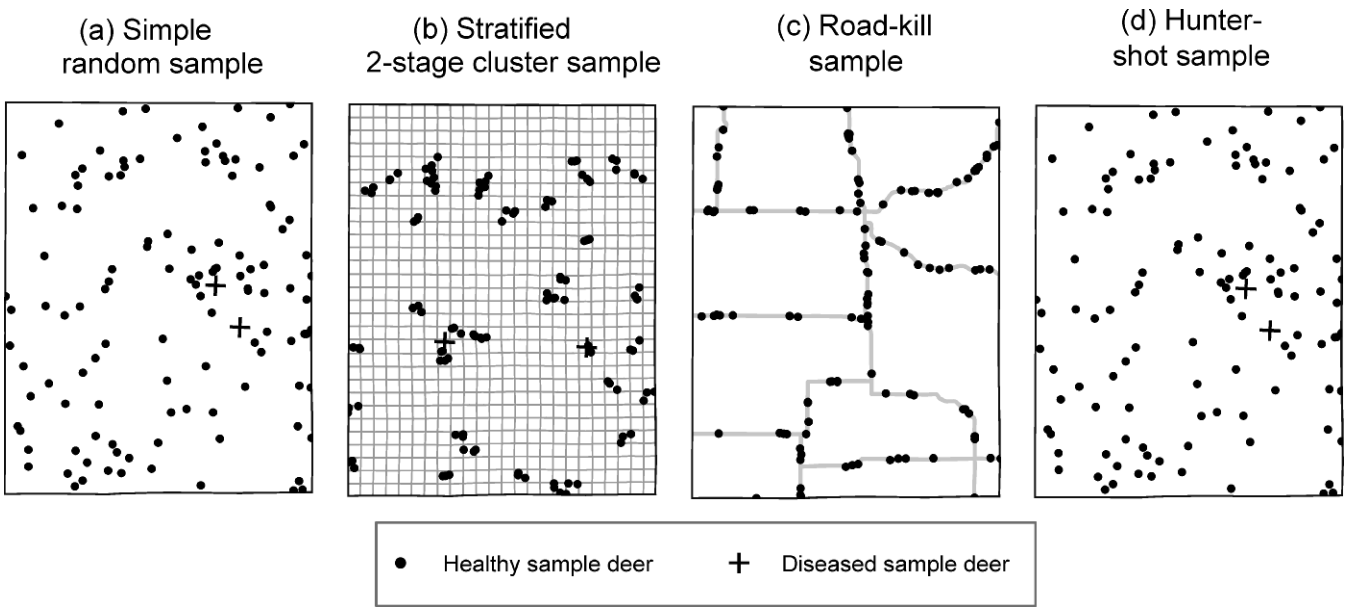

Figure 2. A simulation realization depicting samples selected from a deer population with disease distributed according to a hot spot and spark model for each of 4 deer harvest methods: (a) simple random sample, (b) stratified 2-stage cluster sample for organized sharp shooting, (c) road-kill sample, and (d) hunter-shot sample.

sections with probability proportional to the fraction of forest cover in the section. This approach encouraged oversampling of sections with habitat preferred by deer. In the second stage of the sampling, the sample unit was an animal, and 3 animals were harvested using SRS from each sampled section (Fig. 2b). In practice, this design restricts harvest to the 40 sampled sections, which will tend to have higher animal densities because of the first stage design.

Convenience sample approaches.—Convenience samples can be difficult to simulate in practice. They involve biases that are generally not well understood, and thus some assumptions must be made to generate the samples. For the road-kill convenience sample, we assumed deer in our snapshot were located within 100 m of state and United States highways in the county were available to be sampled (i.e., killed by a vehicle). This buffer represented 3.5% of the area in the county, and it had proportionately less forest (8% in buffer vs. 15% in entire county) and more artificial cover (primarily roadway; 5% in buffer vs. 2% in entire county). On average, 205 deer (3% of the population) were expected to be located within the buffer. We also assumed that animals with the disease, whose symptoms include weakened physical condition and confused behavior, were twice as likely to be hit by vehicles relative to healthy deer (Conner et al. 2000). To implement this approach, we assumed that only animals near a road might be struck by a vehicle, and we selected a sample of deer that fell within the road buffer with probability proportional to the likelihood of being hit. Deer located beyond this road buffer were not sampled by this simulated road-kill mechanism (Fig. 2c). For the hot spot and spark disease model, only migrant diseased animals (i.e., sparks) had the opportunity to fall within road buffers because in our demonstrated application the disease cluster was not intersected by any roads.

To construct a hunter–harvest sample design, we noted that hunters generally spend time where deer are most prevalent. To reflect this situation, we assumed that the number of animals harvested in a habitat type was proportional to the total number of animals located in the habitat extent. This can be approximated by a STS with proportional allocation of the sample size across strata. The STS was a design in which we selected a random sample of animals independently from each stratum (i.e., a subset of the population), which for our simulation was a habitat type. Under this design, the number of animals harvested in each habitat stratum was proportional to total number of animals located in each habitat. Because hunters generally avoid weakened or confused animals, we set the likelihood of harvesting a diseased animal within a stratum to be 0.8 that of a healthy animal. To implement this design, we selected animals from each habitat with probability proportional to the likelihood of being harvested (Fig. 2d).

Estimating surveillance parameters.—We used SRS estimators to calculate estimates of disease prevalence for the benchmark SRS, road-kill and hunter-shot samples. For convenience samples, SRS estimators are often used in the absence of information on selection probabilities for animals. The SRS estimators assume that every possible sample of n animals is equally likely, which is violated by the disease bias conditions simulated in the road-kill and hunter-harvested designs. Let y_k equal 1 if animal k has the disease, and 0 if the animal does not (i.e., is healthy). The SRS estimator for the disease prevalence p is $\hat{p}_{\mathrm{SRS}} = n^{-1} \sum_{k=1}^{n} y_k$.

For 2-stage cluster design used in an organized sharp-shooting campaign, we used the design–unbiased survey estimator, which takes into account the oversampling of forested sections and differing selection probabilities for

Table 1. Probability of detecting disease in an annual sample (i.e., fraction of simulation samples with ≥1 diseased animal) and expected number of annual samples required to detect the presence of disease, for each sample design under each disease model.

	Sampling approach			
	SRS[a]	Stratified 2-stage cluster (sharp shoot)	Road-kill	Hunter-shot
Selection bias	None	None	Towards diseased animals	Against diseased animals
Random disease model				
Probability of detecting ≥1 diseased animal in the sample	0.83	0.84	0.84	0.80
Mean no. of annual samples required to detect ≥1 diseased animal in the sample	1.21	1.20	1.19	1.26
Hot spot and spark disease model				
Probability of detecting ≥1 diseased animal in the sample	0.83	0.85	0.09	0.76
Mean no. of annual samples required to detect ≥1 diseased animal in the sample	1.20	1.17	11.24	1.32

[a] SRS = simple random sampling.

deer across sections. Recall that within stratum h, we drew a sample of $n_h = 4$ sections from the $N_h = 72$ sections from township h, with probability π_{hi} proportional to the fraction of forested area in section i of stratum h. Within this section we randomly sampled $m_{hi} = 3$ animals from the M_{hi} deer present in the section so that the conditional probability of sampling deer k from section i of stratum h was $\pi_{k|hi} = m_{hi}/M_{hi}$. The unconditional probability of selecting a deer from the population was equal to $\pi_{hik} = \pi_{hi}\pi_{k|hi}$ and the estimator for the prevalence under the sharp shooting design is

$$\hat{p}_{Cluster} = \left[\sum_h \sum_i \sum_k \pi_{hik}^{-1}\right]^{-1} \sum_h \sum_i \sum_k \pi_{hik}^{-1} y_{hik},$$

where y_{hik} is the disease status indicator for deer k from section i of stratum h.

We calculated the simulation standard deviation as the standard deviation of the individual prevalence estimates under each of the 8 conditions (2 disease models × 4 sampling approaches). Similarly, we calculated the probability of detecting ≥1 diseased deer in the sample of n animals, θ, for each the sampling approach and disease model as the proportion of samples that had ≥1 diseased deer. Assuming the number of samples that have ≥1 diseased deer follows a binomial distribution, the number of annual samples (yr) required to detect the disease follows a geometric distribution with mean $\tau = 1/\theta$, which we also calculated.

Disease heterogeneity and detection biases.—Note that although we included selection bias in convenience sample designs in the simulation, we did not include other biases. For example, the dependence of CWD prevalence on habitat is not known at this time, and we did not develop a model that reflects varying disease prevalence across habitats. Also, CWD detection is poorer in younger animals (Miller et al. 2000), but we did not incorporate age structure in our simulations, which would be required to simulate this bias. Finally, we assumed perfect detection conditions (i.e., that all animals were completely detectable, diseased animals were always recognized as such when sampled, and healthy animals were always recognized as such when sampled), which implied that the disease test sensitivity and specificity were both 1.

RESULTS

Under the random disease model, the waiting time distribution parameters (i.e., the probability of detecting ≥1 deer in a single sample and the mean no. of samples required to detect ≥1 deer) were approximately the same for the 2 probability sample designs: SRS and the 2-stage cluster sample for sharp shooters (Table 1). If 120 animals were sampled, under these designs, there was at least an 84% chance of observing an infected animal in the first year, and it would take on average ≥1.20 annual samples to detect ≥1 infected animal. Our estimates of these parameters were nearly identical under the hot spot and spark model for both probability sample designs.

Under the random disease distribution model, road-kill waiting time distribution parameters were roughly the same as the probability sample designs (Table 1). Under the hot spot and spark model, the waiting time distribution for the road-kill sample had a very low detection probability (9%) and a long waiting time (>11 yr).

As expected, the hunter-shot convenience samples had a lower detection probability (80% chance of sampling a diseased animal) and higher waiting time (mean of 1.26 samples required to detect a disease-positive animal) than the probability sample designs and the road kill design (Table 1). The detection probability was somewhat lower (76%) and the waiting time longer (1.32 annual samples) for hunter-shot samples than the probability samples under the hot spot and spark disease model.

The mean of the prevalence estimator distribution over simulation replicates was approximately equal to the true prevalence of 1.5% under either disease distribution model for both probability designs when we used the proper estimator (Table 2). The simulation mean of prevalence estimates for the road-kill design depended on the disease model. Under the random disease model, the true prevalence of 1.5% was greatly overestimated (2.2%) by the road-kill

Table 2. Simulation mean and standard deviation of prevalence estimates, for each sample design under each disease model.[a]

	Sampling approach			
	SRS[b]	Stratified 2-stage cluster (Sharp shoot)	Road-kill	Hunter harvest
Selection bias	None	None	Towards diseased animals	Against diseased animals
Estimator	SRS (correct)	Cluster (correct)	SRS (naive)	SRS (naive)
Random disease model				
Mean of prevalence estimates	0.015	0.015	0.022	0.012
SD of prevalence estimates	0.012	0.014	0.015	0.009
Hot spot and spark disease model				
Mean of prevalence estimates	0.015	0.016	0.001	0.012
SD of prevalence estimates	0.011	0.013	0.004	0.010

[a] The true disease prevalence for all conditions is 0.015 (1.5%).
[b] SRS = simple random sampling.

design (Table 2). In contrast, under the hot spot and spark model used in this simulation, the true prevalence was grossly underestimated for the road-kill design. Compared to the true 1.5% prevalence rate, the mean of the prevalence estimates was 0.1%. For the hunter-shot sample, which involves selection bias in favor of healthy animals, the mean of the prevalence estimate is biased downward (1.2%) from the true prevalence for both the random and the hot spot and spark disease distribution models (Table 2).

DISCUSSION

Random Disease Model

Although we generally considered the random disease model to be unrealistic, studying its properties was instructive as a baseline for comparison to more realistic scenarios. The random disease model also demonstrated the impact that convenience sampling selection bias may have on properties of estimates.

Given a low prevalence rate of 1.5% for a randomly distributed rare disease, the choice of sampling design has relatively little effect on the probability of detecting a diseased animal or the waiting time for detecting a diseased animal. Even when we incorporated selection biases that arose from the harvest process (i.e., avoiding or favoring diseased animals) into convenience sample designs, the waiting time distributions did not differ greatly from those of the probability sample designs under the random disease model. For example, we expected the road kill sampling approach with a bias toward diseased deer to result in shorter waiting time and higher detection probability than unbiased probability sampling. But the relative similarity of the results is likely due in part to the fact that we were evaluating parameters for a rare disease with samples drawn from a deer population frame restricted to road buffers (Fig. 2).

However, when we considered the estimated prevalence, selection biases associated with convenience sample did have an effect, which was particularly obvious under the random disease distribution. For example, because we modeled infected animals as being more likely to be killed on the road (Krumm et al. 2005), the average prevalence estimate from the road-kill design was higher than the true prevalence of the population. Note that even if the disease distribution were random across animals, a characteristic of the animals could place them at greater risk for being harvested under the convenience sample from road-kills. Estimates obtained from the road-kills under these conditions would lead wildlife managers to overestimate disease prevalence, although waiting time to detection of the disease would be shortened because of the increased probability of detection.

The opposite problem exists for hunter-shot samples. Given a random disease model and the assumption that hunters avoid infected deer, on average, hunter-shot samples will underestimate the true prevalence of the disease. The implication for the hunter-kill design is that if assumptions of a random disease distribution and simple random sampling are used to calculate a sample size, the probability that a diseased animal will be detected in the sample is likely to be lower than investigators assume. Our results emphasize that using a sample of hunter-shot deer to design a surveillance study may lead to an underestimate of the sample size needed to achieve surveillance goals.

Hot Spot and Spark Model

The hot spot and spark disease model was a more realistic representation of the disease process, and studying its properties yielded insights into the problem of coverage in convenience sampling. Coverage is the degree to which a sampling frame represents the entire study population.

The road-kill design performance depended heavily on the disease model assumptions because the disease cluster was located in interior landscape that did not intersect the road buffer population being sampled. Only the occasional diseased migrants (sparks) had the opportunity to intersect with the buffer, and diseased migrants were very rare in the road buffer population relative to the baseline county-level disease prevalence. Thus, the probability of detection for the road-kill sample was extremely low, average waiting times were very long, and prevalence estimates were severely underestimated. Under this kind of disease scenario, relying solely on road-kill convenience samples would be a very poor basis for surveillance.

In our simulation, the failure of road-kill samples to cover

an important part of the population (i.e., the nucleus of the infection) led to poor surveillance outcomes, but coverage error can work in the other direction. If the disease cluster overlaps significantly with the road buffer, in essence over-representing the infected population relative to the healthy population, a very different picture may emerge. For example, if 20 of the 105 county's diseased animals were located in the road buffer, and an average of 205 animals were located within the buffer, then the average prevalence of the buffer area would be far higher than the 1.5% prevalence for the full county population. The upside of over-representing infected individuals is that it is nearly certain that the disease would be detected in the first year's sample of 120 animals, which could then trigger a more rigorous sampling approach for prevalence estimation.

Under our simulation, there was no coverage error associated with hunter-shot samples. Because of this, the bias in surveillance parameters for hunter-shot samples was very similar under both disease-distribution models. However, if the population of hunters restricts their hunting to areas that omitted (or alternatively, included) an infection cluster, coverage error would have an effect on estimated surveillance parameters.

Partial Implementation of Probability Designs

Field scientists often turn to convenience sampling because they believe that probability sampling is impractical, and it is often impossible to implement a random selection mechanism for individual population elements (e.g., animals). However, in practice there are several methods of partially implementing probability sampling to mitigate the biases of convenience sampling. For example, 2-stage cluster sample designs involve randomly selecting small areas that can reasonably be traversed for harvesting a small sample of animals via sharp shooting. It is generally straightforward to draw the first stage sample of land areas using a probability sample design. Because it is usually impossible to randomly select animals in the sampled land areas, the second stage sample might consist of convenience sample of deer opportunistically shot in the land area. In doing so, scientists can remove bias by randomly selecting land areas for sharp-shooting, and can mitigate biases in the convenience sample by training the sharp-shooters to avoid biases in selection of animals.

Field data collection is generally more efficient if we collect samples from areas that are likely to yield more information on the disease. Probability sampling schemes can be applied that lead to probability-based oversampling (i.e., sampling at a higher rate) of more informative first-stage sample areas (e.g., forested areas for deer), so that field time would be more efficiently used. This is a form of unequal probability sampling called sampling with probability proportional to size (PPS). The size measure is a measure of the importance of a sample unit with respect to providing information on the primary variables in a survey. If the size measure is related to the likelihood that the disease is present in the sample unit, then PPS leads to higher precision of the estimates (Lohr 1999). In our simulation, the size measure was the proportion of forest in the section, which would usually lead to samples of clusters with more animals in them. Because of simulation assumptions, forest cover was either not related (random disease model) or only slightly related (hot spot and spark model) to disease prevalence, and thus we did not expect substantive differences in the precision in the results, even though cluster samples usually have poorer precision. Nonetheless, if information is available about factors that indicate the presence of the disease (Joly et al. 2006), then these factors can be used with PPS to improve the operational and statistical efficiency of cluster samples.

If oversampling is used, scientists must construct weighted estimates to correct for the higher rate of sampling of some types of clusters relative to other types of clusters. If weights are not used, (e.g., if investigators use SRS formulas instead of the proper 2-stage cluster sample estimator) estimates and standard errors will be biased. This principle is analogous to making sure that the correct analysis of variance is applied for the experimental design and that the correct error terms are used to test treatment effects. References on methods (Lohr 1999, Thompson 2002) and software packages (e.g., Stata [Stata Corp LP, College Station, TX], SAS [SAS Institute, Cary, NC], R [R Project Network, http://www.R-project.org]) to calculate proper weighted means and totals are readily available.

Stratification is used to improve the likelihood that the sample provides a good representation of the population. In our simulation, we used township strata to ensure that all parts of the county would be represented in the sample. This is like blocking in experimental designs in that independent samples are taken from each partition of a population. In survey sampling, stratification provides more precise estimators if the factor used to partition the study area into sampling strata is related to disease prevalence (i.e., the partition results in some strata having notably higher disease prevalence than others). When using convenience samples, stratification should be and often is used to disperse the sample throughout the study area.

Use of Simulation to Evaluate Surveillance Designs

We developed a framework of simulating healthy and diseased deer distributions using geographic information from real mixed forest and agricultural landscapes and demonstrated the value of thinking about surveillance in a formal sampling context. Although we applied simple assumptions about the ecological properties and the disease sampling, researchers could add more detail as data are collected to refine understanding about disease distribution and transmission. For example, we could simulate density-dependent disease transmission (Joly et al. 2006) to study its effect on relative performance of alternative sampling designs. Our framework also could be used to simulate regions where habitat is more continuous and deer are very dense, and it could be very profitably applied to larger areas like multiple counties or deer management units in a state.

Although convenience sampling will likely be needed, its bias can be carefully considered in a simulation environment

so that the best possible field sampling method can be constructed. By starting with relevant ecological data and theory, an artificial population of healthy and diseased deer can be distributed on the landscape of interest, and the statistical properties of detection, prevalence and waiting time parameters estimated under alternative designs can be studied. Also, it is critical to identify the objective of the surveillance activity (e.g., detection vs. prevalence estimation) as part of the process of identifying effective candidate and benchmark designs.

MANAGEMENT IMPLICATIONS

Convenience samples of animals used to collect observational data for disease surveillance are often implicitly or explicitly assumed to be representative of a population. However, sampling biases invariably occur, and the inability to quantify the probability of obtaining a sample unit in a convenience sample makes it impossible to develop a statistically valid estimate of surveillance parameters. We recommend that wildlife scientists fully appreciate and use this fundamental connection between probability-based and observational data. Investigators can use simulations to model knowledge about the wildlife population, disease epidemiology, and the inherent biases in convenience samples. Investigators can also use simulations to evaluate alternative sampling approaches in relation to probability designs. In addition, although it is impractical to fully apply probability sampling, it is often possible to invoke probability sampling devices (e.g., stratification, cluster sampling, sampling with probability proportional to an importance factor) for some part of the sampling process to reduce the impact of bias generated by convenience sampling.

ACKNOWLEDGMENTS

Our research was funded in part by the United States Geological Survey National Wildlife Health Center, and it is a contribution of United States Department of Agriculture Multistate Project NC-1005. We thank H. Butler for producing the figures, and 2 anonymous reviewers for helpful comments.

LITERATURE CITED

Beringer, J., L. P. Hansen, J. J. Millspaugh, and T. Meyer. 2003. A statewide surveillance effort for detecting chronic wasting disease in wild white-tailed deer in Missouri. Wildlife Society Bulletin 31:873–879.

Conner, M. M., C. W. McCarty, and M. W. Miller. 2000. Detection bias in harvest-based estimates of chronic wasting disease prevalence in mule deer. Journal of Wildlife Diseases 36:691–699.

Courchamp, F., L. Say, and D. Pontier. 2000. Detection, identification, and correction of a bias in an epidemiological study. Journal of Wildlife Disease 36:71–78.

Diefenbach, D. R., C. S. Rosenberry, and R.C. Boyd. 2004. Efficacy of detecting chronic wasting disease via sampling hunter-killed white-tailed deer. Wildlife Society Bulletin 32:267–272.

Doherr, M. G., and L. Audige. 2001. Monitoring and surveillance for rare health-related events: a review from the veterinary perspective. Philosophical Transactions of the Royal Society of London B 356:1097–1106.

Edwards, D. 1998. Issues and themes for natural resources trend and change detection. Ecological Applications 8:323–325.

Eidson, M., N. Komar, F. Sorhage, R. Nelson, T. Talbot, F. Mostashari, and R. McLean. 1999. Crow deaths as a sentinel surveillance system for West Nile Virus in the northeastern United States. Emerging Infectious Diseases 7:615–620.

Huang, L. 2005. Evaluating sample designs to detect chronic wasting disease in deer populations. Masters Creative Component. Iowa State University, Ames, USA.

Joly, D. O., C. A. Ribic, J. A. Langenberg, K. Beheler, C. A. Batha, B. J. Dhuey, R. E. Rolley, G. Bartelt, T. R. Van Deelen, and M. D. Samuel. 2003. Chronic wasting disease in free-ranging Wisconsin white-tailed deer. Emerging Infectious Diseases 9:599–601.

Joly, D. O., M. D. Samuel, J. A. Langenberg, J. A. Blanchong, C. A. Batha, R. E. Rolley, D. P. Keane, and C. A. Ribic. 2006. Spatial epidemiology of chronic wasting disease in Wisconsin white-tailed deer. Journal of Wildlife Diseases 42:578–588.

Krumm, C. E., M. M. Conner, and M. W. Miller. 2005. Relative vulnerability of chronic wasting disease infected mule deer to vehicle collisions. Journal of Wildlife Diseases 41:503–511.

Lohr, S. 1999. Sampling: design and analysis. Brooks/Cole, Pacific Grove, California, USA.

Miller, M. W., and M. M. Conner. 2005. Epidemiology of chronic wasting disease in free-ranging mule deer: spatial, temporal, and demographic influences on observed prevalence. Journal of Wildlife Diseases 41:275–290.

Miller, M. W., E. S. Williams, C. W. McCarty, T. R. Spraker, T. J. Kreeger, C. T. Larsen, and E. T. Thorne. 2000. Epizootiology of chronic wasting disease in free-ranging cervids in Colorado and Wyoming. Journal of Wildlife Diseases 36:676–690.

Mörner, T., D. L. Obendorf, M. Artois, and M. H. Woodford. 2002. Surveillance and monitoring of wildlife diseases. Review Scientifique et Technique Office Internationale des Epizooties 21:67–76.

Nusser, S. M., F. J. Breidt, and W. A. Fuller. 1998. Design and estimation for investigating the dynamics of natural resources. Ecological Applications 8:234–245.

Nusser, S. M., and E. E. Klaas. 2003. Survey methods for assessing land cover map accuracy. Environmental and Ecological Statistics 10:309–331.

Olsen, A. R., J. Sedransk, C. Gotway, W. Liggett, S. Rathbun, K. Reckhow, L. Young, and D. Edwards. 1999. Statistical issues for monitoring ecological and natural resources in the United States. Environmental Monitoring and Assessment 54:1–45.

Otis, D. L. 2001. Quantitative training of wildlife graduate students. Wildlife Society Bulletin 29:1043–1049.

Samuel, M. D., D. O. Joly, M. A. Wild, S. D. Wright, D. L. Otis, R. W. Werge, and M. W. Miller. 2003. Surveillance strategies for detecting chronic wasting disease in free-ranging deer and elk: results of a CWD surveillance workshop, 10–12 December 2002. United States Geological Survey National Wildlife Health Center, Madison, Wisconsin, USA.

Schmitt, S. M., S. D. Fitzgerald, T. M. Cooley, C. S. Bruning-Fann, L. Sullivan, D. Berry, T. Carlson, R. B. Minnis, J. B. Payeur, and J. Sikarskie. 1997. Bovine tuberculosis in free-ranging white-tailed deer from Michigan. Journal of Wildlife Diseases 33:749–758.

Shaffer, T. L., and D. H. Johnson. 2008. Ways of learning: observational studies versus experiments. Journal of Wildlife Management 72:4–13.

Taylor, C. E., D. L. Otis, H. E. Hill, Jr., and C. R. Ruth. 2000. Design and evaluation of mail surveys to estimate deer harvest parameters. Wildlife Society Bulletin 28:717–723.

Thompson, S. K. 2002. Sampling. John Wiley and Sons, New York, New York, USA.

U.S. Geological Survey. 1992. The National Land Cover Data 1992. <http://eros.usgs.gov/products/landcover/nlcd.html>. Accessed 13 Nov 2007.

Vennette, R. C., R. D. Moon, and W. D. Hutchinson. 2002. Strategies and statistics of sampling for rare individuals. Annual Review of Entomology 47:143–174.

Ver Hoef, J. 2002. Sampling and geostatistics for spatial data. Ecoscience 9: 152–161.

Williams, E. S., M. W. Miller, T. J. Kreeger, R. H. Kahn, and E. T. Thorne. 2002. Chronic wasting disease of deer and elk: a review with recommendations for management. Journal of Wildlife Management 66: 551–563.

Associate Editor: Shaffer and Johnson.