AN EFFICIENT STRATEGY TO ESTIMATE INTENSITY AND PREVALENCE: SAMPLING METACERCARIAE IN FISHES

Author(s): Jenny C. Shaw, Leopoldina Aguirre-Macedo, Kevin D. Lafferty

AN EFFICIENT STRATEGY TO ESTIMATE INTENSITY AND PREVALENCE: SAMPLING METACERCARIAE IN FISHES

Jenny C. Shaw, Leopoldina Aguirre-Macedo*, **and Kevin D. Lafferty**†

Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, California 93106. *e-mail: shaw@lifesci.ucsb.edu*

ABSTRACT: Accurate estimates of population-level parameters of parasites, such as prevalence and mean intensity, require large sample sizes. The processing of such samples becomes an overwhelming task when parasites are abundant, as with trematode metacercariae in fishes. In the present study, a subsampling method reduced processing time while maintaining an accurate estimation of metacercariae prevalence and intensity across 3 trematode species and 2 fish species. By double sampling, we generated regression models to predict total intensity from a combination of subsamples. The key to this approach lies in choosing the best strategy from a large number of potential subsampling routines. We selected the most efficient routine by weighing the costs and benefits of each. This approach, however, could not provide an estimate of parasite abundance when no parasites occurred in the initial subsample. To estimate prevalence accurately, our subsampling algorithm prescribed an additional sampling sequence using a new, optimal regression model. In addition, we optimized the technique to measure three parasite species infecting a single host simultaneously. This efficient subsampling procedure decreased the overall processing time per host by up to 91% while obtaining accurate ($R^2 > 0.8$) estimates for both prevalence and intensity.

Individuals differ greatly in the number of parasites they harbor. When parasites are numerous and widely distributed within the host, comprehensive counts can be extremely tedious, even after a few hosts. After counting tens of thousands of metacercariae from hundreds of fishes, we wondered if such counts could be obtained with less effort.

Prevalence (proportion infected), mean intensity (parasites per infected individual), and mean abundance (parasites per individuals examined) are 3 common population-level descriptors of parasite abundance (Bush et al., 1997). The accuracy and precision of these estimates increases with the number of hosts examined and with the quality of counts within each host. Such comprehensive examinations can be extremely time-consuming. One approach is to sample until the estimate converges on a stable mean within a predetermined confidence limit as determined by bootstrapping, jackknifing, or parametric means. A way to further reduce the time spent per host is to examine specific tissues or organs (subsamples) rather than the entire host. An estimate of total intensity can then be extrapolated from the subsample. This technique sacrifices the accuracy of the count for individual hosts; however, it ultimately may increase the accuracy of population-level parameters by greatly increasing the number of hosts that can be sampled in a given time period.

Estuarine fish that serve as the second intermediate host to trematode parasites provide an ideal system in which to apply a subsampling strategy. Several of these trematodes infect more than one tissue or organ in the fish; furthermore, these site generalists may be distributed throughout the host, with mean intensities of several hundred per fish. Difficulties in evaluating such high intensities are complicated further by the large sample of hosts needed for accurate measures of prevalence and mean intensity. To increase the efficiency of sample processing and, ultimately, the accuracy of population-level estimates, we developed a subsampling approach

in the context of a large-scale, 5-yr project that examines the role of trematodes in estuaries throughout Southern California and Baja California.

The simplest subsampling strategy employs ratios (i.e., the estimate of total intensity:subsample intensity). For parasitologists, this usually entails sampling half the bilaterally symmetrical organs; the estimate of total intensity would be twice the subsample intensity. Because there can be a left-side or a right-side bias in the distribution of parasites (Thiemann and Wassersug, 2000), however, it must be demonstrated first that intensity does not vary significantly between the left and right halves (Graczyk, 1991; Marcogliese et al., 2001). If infected organs are not bilaterally symmetrical, then one can build a linear regression model that predicts the total intensity (I) from 1 or a set of tissues in a subsample (S). Building such a regression model requires a sample where both the total count (I) and the subsample count (S) are obtained from each host. The process for this approach is called double sampling (Cochran, 1977). Here, the subsample, S, serves as the independent variable, and a least-squares linear regression equation (slope [b] and intercept [a]) predicts the relationship between I and S in the standard form

$$\hat{I} = bS + a \qquad (1)$$

where $\hat{I}$ is the predicted value of I for each S. Equation (1), with values of b and a obtained from the initial sample, S_1 and I_1, may permit the estimation of intensity from subsequent sampling efforts or host populations with the same values of b and a, where S_2 is the host subsample from a second population with unknown I_2. The estimation of total intensity in a second sample ($\hat{I}_2$) is based on the assumption that the relationship between total intensity (I) and the intensity observed in the subsample (S) is the same for both the initial and subsequent samples; that is, the regression equation does not differ among samples ($\hat{I}_2 \approx b_1 S_2 + a_1$). The larger the proportion of variation explained by the model (as expressed by the coefficient of determination, R^2), then the greater the accuracy of the prediction of intensity. Subsampling will be an efficient method of generating an accurate estimate of mean intensity if

$$R^2 > (4c_t c_s)/(c_t + c_s)^2 \qquad (2)$$

where c_t is the average time to process an entire host and c_s is

* Centro de Investigación y Estudios Avanzados del IPN Unidad Mérida, Antigua Carretera a Progreso Km 6, Cordemex, Mérida, Yucatán, C.P. 97310, Mexico.

† USGS, Western Ecological Research Center, c/o Marine Science Institute, University of California, Santa Barbara, California 93106.

the average time to process a subsample (Cochran 1977, p. 341). In addition, this approach assumes that the distribution of metacercariae among sites within a host does not vary with intensity (we verify this assumption below).

Most double sampling programs choose a subsampling strategy in advance based on practical reasons or expert opinion. The strategy is then evaluated (as described above) to determine if it is more efficient than complete sampling. If several organs are infected by a parasite, however, then a subsample could consist of any one organ or combination thereof. Therefore, a large number of potential subsampling strategies will exist, because the number of possible combinations (including sampling all organs) from a list of n organs is $2^n - 1$. Subsequently, the most optimal subsampling strategy may not be readily apparent. Here, we provide a method to consider trade-offs in sampling effort and accuracy. This general approach is flexible enough to be widely applied.

MATERIALS AND METHODS

Animal collection and parasite identification

Seventy-three *Clevelandia ios* and 116 *Fundulus parvipinnis* were seined from Estero de Punta Banda (Ensenada, Baja California) in October 2002 and November 2002. Individuals were bagged according to species, killed with dissolved CO_2, and transported on ice. Fishes were refrigerated at 4 C for immediate dissection; those not examined within 48 hr were frozen and stored at -20 C. Dissections were performed with the aid of a dissecting microscope. External structures and internal organs were removed, squashed, and examined with a compound microscope for the counting and identification of parasites. Because initial investigations did not indicate a left- or right-side bias in infection, we examined the left side of paired structures (e.g., eyes, gills, pectoral fins, and lateral musculature), and we doubled the count to estimate the total intensity for both body parts. We found a high association (least-squares regression, $R^2 = 0.94$, $P < 0.001$, n = 34 body parts) between predicted ($2 \times$ left) and observed (left + right) counts. This bilateral sampling was done to simplify our approach and reduce effort. However, it is important to acknowledge that simply by doubling counts of the left side of bilateral organs, our data underestimate the error in the estimates, because the actual total count of parasites from a bilateral organ equals 2 times the left side plus a random error term, and we have assumed this random error term to be 0.

To determine the efficiency of further subsampling, we assessed all species of metacercariae infecting both fish hosts. Several species occurred in both fishes. Subsampling seemed to be inappropriate for 4 species, because they inhabited organs that were either difficult to subdivide (*Phocitremoides ovale* and *Renicola buchanani* in both fishes and *Euhaplorchis californiensis* in *F. parvipinnis*) or had such low intensity that total counts were not cumbersome (*Mesostephanus appendiculatus* in both fishes). Three tissue generalists appeared to be promising for further subsampling. *Stictodora hancocki* and an unknown cyathocotylid species were found in both fish hosts. In addition, *Phocitremoides* sp. occurred in *C. ios*. For all 3 parasites, intensity did not alter significantly the distribution of metacercariae among organs, suggesting that a linear regression model was appropriate for predicting intensity from a subsample; that is, no significant correlations were found between the intensity and the proportion of metacercariae in a fish that occurred in a particular organ when compared with a Bonferroni-adjusted P-value for multiple comparisons. All metacercariae were identified with a compound microscope following descriptions by Martin (1950a, 1950b) and Yamaguti (1971).

Subsample strategy evaluation

For a given fish species, the number of potential subsampling strategies was calculated by a permutation method for all possible combinations of organs, tissues, or both in which a particular parasite species was found. The 12 organs we sampled from *C. ios* provided 4,095 potential subsampling strategies, and the 13 organs we sampled from

F. parvipinnis provided 8,191 potential subsampling strategies. We evaluated the effort required for each strategy by estimating the processing, handling, and counting time needed to examine and analyze an average infected subsample (excluding uninfected subsamples). Table I lists the average time required to process each organ for *C. ios* and *F. parvipinnis*. We defined the accuracy, or predictive power, of each possible subsampling strategy as the percentage of the variation in the total count explained by the subsample count (R^2) for parasites. Our approach was complicated by the fact that each fish species had more than 1 parasite species (often 2 or more in a particular organ). We wished to develop a procedure that would sample all parasite species efficiently. For this reason, the effort for a particular subsampling strategy included the combined effort to sample each parasite species, and the predictive power was expressed as the average R^2 for each parasite species.

We placed 2 additional constraints on our choice for the most optimal subsampling strategy. First, we rejected those subsampling strategies in which less than 60% of the fish had subsamples infected with each parasite species to avoid choosing a strategy that would yield too many uninfected subsamples. Second, we placed a higher priority on accuracy over effort reduction, and we considered only strategies with a minimum R^2 of 0.80. We expressed the efficiency (E) of a subsampling strategy by rearranging Equation 2 so that $E = R^2(c_t + c_s)^2/(4c_tc_s)$, where the subsampling was more efficient than complete sampling if $E > 1$. Assessing efficiency for up to 8,192 combinations created a computational challenge. As a result, we programmed a spreadsheet to calculate the efficiency of subsampling a randomly selected combination of organs. Then, we generated more than 50,000 iterations to ensure the likelihood of evaluating all possible combinations. The subsampling strategy with the highest efficiency was chosen for each host.

For the most efficient subsampling strategy, we calculated the slope (b) and intercept (a) as described previously for each parasite species in each fish species. In addition, we calculated the standard error (SE) of each estimate of intensity (Sokal and Rohlf, 1981) as

$$\text{SE of } \hat{I}_2 = \sqrt{\frac{SSEI_1}{n_1 - 2}\left[1 + \frac{1}{n_1} + \frac{(S_2 - \bar{S}_1)^2}{SSS_1}\right]}$$

where $SSEI_1$ is the sum of squares error of I_1, or $\Sigma(I_1 - \hat{I}_1)^2$ and SSS_1 is the sum of squares of S_1, or $\Sigma(S_1 - \bar{S}_1)^2$. To estimate the standard error from a sample of several fish, we calculated the mean and variance of $\hat{I}_2$ (where $\hat{I}_2 = $ the predicted intensity from a subsample of a subsequent host population or sampling effort), which approximate the actual mean and variance of I_2 (where $I_2 = $ the actual intensity from a subsequent host population or sampling effort).

The calculation of slope, intercept, and standard error allowed us to estimate accurately the intensity from each fish sampled using the most efficient subsampling strategy. This, however, left us without an estimate for individual fish in which the subsampled organs were uninfected. Rather than assuming based on an uninfected subsample that an entire fish had no parasites, we determined the most efficient sequence for adding individual organs to the subsample until an infection for each parasite was observed. The relative benefit of adding a particular organ was expressed as the ratio of the prevalence divided by the intensity for a particular organ averaged across parasite species. Again, a regression equation was fitted to estimate the intensity of a particular host–parasite combination with every new potential subsampling strategy (the addition of new organs to the list of independent variables necessitated a different equation). More importantly, this sequential sampling approach allowed a precise determination of prevalence by requiring the processor to sample all the organs of an uninfected fish. Finally, we expressed the relative effort of subsampling as a percentage of the total time required to process an average infected fish.

Simulation model

We assessed our approach by a simulation model using, for simplicity, the *F. parvipinnis* samples. We randomly selected (with replacement) 2 groups of 50 fish from the original 116 fish. From the first group, we calculated regression equations between the subsamples and originals according to our protocol. We then applied these regression equations to subsamples from the second group to estimate the total intensity in the second group. We compared the mean and standard deviation of the estimate of the second sample with the actual mean and standard deviation of the second sample. This was repeated 1,000

TABLE I. The average time (min) required to process whole organs (c_t). The second number for bilateral organs (in parentheses) represents processing time for only the left side (c_s). The entire pelvic fin was sampled for the arrow goby (*Clevelandia ios*).

Fish sp.	Parts sampled													
	Anal fin	Body cavity	Brain/head musculature	Caudal fin	Dorsal fin	Digestive tract	Eye*	Gills*	Liver	Body musculature*	Orobranch- ial cavity	Pectoral fins*	Pelvic fins*	Scales*
Clevelandia ios	1.7	6.5	9.9	5.3	1.2	4.7	3.3 (1.6)	45.9 (22.9)	0.6	39 (19.5)	5.9	4.9 (2.5)	4	1 (0.5)
Fundulus parvipinnis	2.1	2.5	—	10.2	3.1	4.2	9.4 (4.7)	51 (25.5)	1.6	37.2 (18.6)	—	9 (4.5)	3.3 (1.7)	1.5 (0.7)

* Bilateral body part (sampled on the left side only).

times. From the simulations, we calculated the number of times the estimated mean and 95% prediction limits overlapped the actual value. We also assessed the degree of association (proportion of the explained variation) between estimated and actual values.

RESULTS

In a typically infected arrow goby (*C. ios*) from Estero de Punta Banda, a total count of the metacercariae of *S. hancocki*, *P. ovale*, and the unknown cyathocotylid species took 134 min and involved sampling the eyes, gills, orobranchial cavity, liver, digestive tract, body cavity, pectoral fins, pelvic fin, dorsal fin, anal fin, caudal fin, brain/head musculature, and body musculature. Sampling only the left side of the bilateral organs (eyes, gills, pectoral fins, and body musculature) reduced our processing time to 87 min per fish. The most efficient subsampling strategy consisted of examining only the body cavity, brain/head musculature, left side of the body musculature, and left pectoral fin; this approach took 39 min for an average infected fish (or 29% of the total processing time). The average R^2 across all 3 parasite species was 0.92 ($P < 0.001$). This was an efficient sampling effort ($E = 1.3$). Table II presents slopes, intercepts, and R^2 values for the association between subsample and total counts for each parasite species. For subsamples that were uninfected by 1 or more parasite species (25% of samples), the sequence of additional organs was, in order of efficiency, the pelvic fin, left eye, right gills, digestive tract, orobranchial cavity, caudal fin, dorsal fin, anal fin, and liver.

In a typically infected killifish (*F. parvipinnis*) from Estero de Punta Banda, a total count of the metacercariae of *S. hancocki* and the unknown cyathocotylid species took 133 min and involved sampling the eyes, gills, liver, digestive tract, body cavity, pectoral fins, pelvic fins, dorsal fin, anal fin, caudal fin, body musculature, and scales. Sampling only the left-side bilateral organs (eyes, gills, pectoral and pelvic fins, and body musculature) reduced sampling effort to 80 min per fish. The most efficient subsample routine consisted of the left eye, the body cavity, the left pectoral fins, and the scales on the left side of the fish; this approach took 12 min for an average infected fish (or 9% of the total processing time). The average R^2 across parasite species for this approach was 0.81 ($P < 0.001$). The substantial reduction in time made this a very efficient sampling strategy ($E = 2.6$). Table III presents slopes, intercepts, and R^2 for the association between subsample and total counts for each parasite species. For the 40% of subsamples that were uninfected by 1 or more parasite species, the additional infection sites to examine were, in order of efficiency, the left body musculature, left gills, caudal fin, digestive tract, liver, left pelvic fin, dorsal fin, and anal fin.

The resampling of our *F. parvipinnis* data indicated a correspondence between the actual and the predicted values, suggesting that the approach was useful for both species of parasites. The 95% prediction limits for the estimation of *S. hancocki* intensity (average of means = 34.9) overlapped the actual value (average of means = 34.7) 99.6% of the time. In addition, the association between the estimates and the actual values had an R^2 of 0.70 ($P < 0.001$). The estimated standard deviation (average of means = 43.9, n = 1,000) was correlated with the actual standard deviation (average of standard deviations = 46.2; $R^2 = 0.41$, $P < 0.001$). The 95% confidence limits for the estimation of intensity for the unknown cyathocotylid spe-

TABLE II. Effort and regression statistics of subsampling for 3 species of trematode metacercariae in 73 arrow gobies (*Clevelandia ios*). Each subsampling effort sampled all 3 species. Columns represent a range of subsampling procedures from a minimum of 4 body parts (with an effort of 29%) to the last column, "+ Liver," which represents sampling all organs. Statistics include the least-squares linear regression slope (b), the intercept (a), the coefficient of determination (R^2), the size of the initial sample (n_1), the average intensity in the subsample ($\bar{S}_1$), the sum of squares error of intensity ($SSEI_1$), and the sum of squares of the subsample (SSS_1). The last 4 variables are required for determining the standard error of the estimate of intensity from a subsample obtained from a single fish. All *P*-values were less than 0.001. The final column does not show a 100% effort because of the time saved by sampling only half the bilateral organs.

Trematode sp./statistic	Body cavity, brain/head musculature, body musculature*, and pectoral fin*	+ Pelvic fin*	+ Eye*	+ Gills*	+ Digestive tract	+ Orobranchial cavity	+ Caudal fin	+ Dorsal fin	+ Anal fin	+ Liver (all organs)
Stictodora hancocki										
b	1.08	1.08	1.08	1.08	1.07	1	1	1	1	1
a	91	90	86	52	43	1	1	1	1	0
R^2	0.90	0.9	0.91	0.92	0.93	1	1	1	1	1
% Infected	97%	97%	97%	97%	99%	99%	99%	99%	99%	99%
n_1	71	71	71	71	72	72	72	72	72	72
$\bar{S}_1$	269	269	273	305	312	375	375	375	375	375
$SSEI_1$	970,616	960,756	902,374	476,279	340,625	234	234	234	234	0
SSS_1	9,287,224	9,279,018	9,345,301	9,524,444	10,007,049	12,130,607	12,130,607	12,130,607	12,130,607	12,190,834
Unknown cyathocotylid sp.										
b	1.09	1.08	1.07	1.02	1.01	1.01	1	1	1	1
a	14	13	12	1	0	0	0	0	0	0
R^2	0.87	0.88	0.88	1	1	1	1	1	1	1
% Infected	93%	95%	95%	96%	96%	96%	96%	96%	96%	96%
n_1	68	69	69	70	70	70	70	70	70	70
$\bar{S}_1$	117.8	118.1	119	135	136	137	139	139	139	139
$SSEI_1$	52,310	43,610	39,704	1,224	618	246	0	0	0	0
SSS_1	1,586,887	163,668	1,648,961	2,105,722	2,121,195	2,137,061	2,175,051	2,175,051	2,175,051	2,175,051
Phocitremoides sp.										
b	1.49	1.28	1.28	1.18	1.18	1.18	1.06	1.04	1	1
a	10	15	15	12	12	9	1	−1	0	0
R^2	0.99	0.99	0.99	0.99	0.99	0.99	0.99	1	1	1
% Infected	75%	79%	79%	82%	82%	82%	85%	88%	89%	89%
n_1	55	58	58	60	60	60	62	64	65	65
$\bar{S}_1$	115.1	123.6	124	132	132	134	152	153	156	156
$SSEI_1$	1,014,671	47,949	479,499	250,152	250,152	241,878	32,164	13,631	0	0
SSS_1	3,166,421	4,387,762	4,387,762	5,161,921	5,161,921	5,166,898	6,491,549	6,792,766	7,402,042	7,402,042
Total effort	29%	32%	33%	50%	53%	58%	62%	63%	64%	65%

* Bilateral body part (sampled on the left side only).

TABLE III. The effort and predictive power of subsampling for 2 species of trematode metacercariae in 116 California killifish (*Fundulus parvipinnis*). Each subsampling effort sampled both organs. Columns represent a range of subsampling procedures from a minimum of 4 body parts (with an effort of 29%) to the last column, "+ Anal fin," which represents sampling all organs. Statistics include the least-squares linear regression slope (b), the intercept (a), the coefficient of determination (R^2), the size of the initial sample (n_1), the average intensity in the subsample ($\bar{S}_1$), the sum of squares error of intensity ($SSEI_1$), and the sum of squares of the subsample (SSS_1). The last 4 variables are required for determining the standard error of the estimate of intensity from a subsample obtained from a single fish. All P-values were less than 0.001. The final column does not show a 100% effort because of the time saved by sampling only half the bilateral organs.

					Parts sampled				
Trematode sp./statistic	Body cavity, eye*, pectoral fin*, scales*	+ Body musculature*	+ Gills*	+ Caudal fin*	+ Digestive tract	+ Liver	+ Pelvic fin*	+ Dorsal fin	+ Anal fin (all organs)
Stictodora hancocki									
b	1.18	1.07	1.08	1.08	0.99	1	1	1	1
a	11	8	0	0	1	0	0	0	0
R^2	0.81	0.86	0.92	0.92	1	1	1	1	1
% Infected	61%	67%	73%	74%	78%	78%	78%	78%	78%
n_1	71	78	85	86	91	91	91	91	91
$\bar{S}_1$	33.7	37.4	43.7	43.3	43.6	44.3	44.3	44.3	44.3
$SSEI_1$	110,644	148,640	170,162	172,119	226,566	224,649	224,649	224,649	224,649
SSS_1	23,696	9,247	2,083	2,021	65	0	0	0	0
Unknown cyathocotylid sp.									
b	2.8	1.21	1.08	1.12	1.12	1.08	1.05	1	1
a	318	42	0	0	0	0	0	0	0
R^2	0.83	0.93	0.91	0.98	0.98	0.99	0.99	1	1
% Infected	92%	99%	100%	100%	100%	100%	100%	100%	100%
n_1	107	115	116	116	116	116	116	116	116
$\bar{S}_1$	224.1	653.6	728.7	825.3	825.3	825.3	870.0	889.2	925.5
$SSEI_1$	22,627,110	103,113,436	136,505,997	170,426,522	170,426,522	170,426,522	185,351,114	195,075,786	220,281,321
SSS_1	128,703,492	26,072,217	10,676,515	3,747,416	3,747,416	3,747,416	1,694,865	853,002	0
Total effort	9%	23%	42%	50%	53%	54%	56%	58%	59%

* Bilateral body part (sampled on the left side only).

TABLE IV. Summary of steps for developing a subsampling approach by double sampling.

Step	Procedure
1	Collect an initial sample of hosts.
2	Count the parasites in a host, recording the site(s) of infection for each.
3	Determine the effort (time) for each site of infection.
4	Determine the number of potential subsampling strategies.
5	Build a predictive model for each strategy.
6	Calculate effort (time cost) and accuracy (R^2) for each potential strategy.
7	Choose the most optimal subsampling strategy.
8	Be sure that subsampling is more efficient than complete sampling.

cies (average of means = 1,058) overlapped the actual value (average = 1,028) 97.4% of the time, with an association of R^2 = 0.61 ($P < 0.001$). The estimated standard deviation (average = 1,400) was correlated with the actual standard deviation (average = 1,502; R^2 = 0.56, $P < 0.001$).

DISCUSSION

Double sampling enables parasitologists to increase the efficiency of subsampling host populations without greatly compromising the accuracy and precision of prevalence and mean intensity estimates. The value of subsampling increases with the savings in cost (i.e., time required to process a subsample) and/or an improvement in accuracy of the estimation. These 2 factors will determine whether total sampling or subsampling gives the best estimate of mean intensity with a particular effort. Our results show that subsampling reduced substantially the effort required to process a sample of fish and to estimate the prevalence and intensity of tissue-generalist metacercariae. For future samples, this subsampling technique can reduce processing time by up to 71% for arrow gobies and by up to 91% for killifish. Despite the reduction in effort, subsample counts still explained a large proportion of the variation in the total count for both fish hosts (arrow goby, 92%; killifish, 81%). Table IV delineates our steps for developing a subsampling technique.

Our approach effectively decreased the effort for estimating mean intensity of a host population; however, it may not be suitably accurate for the estimation of intensity in individual hosts (a matter that depends on R^2 and the residual from the regression line for that host). The intercept (a) often is greater than zero, meaning the linear regression approach would estimate that all uninfected subsamples have an intensity = a. An alternative is to assume that $a = 0$ (forcing the association through the origin), essentially using a ratio approach. This option has 2 drawbacks. First, it falsely concludes that if the subsample has no parasites, then the host is always uninfected. Second, forcing the association through the origin also reduces the predictive power (decreases the R^2) by constraining the best fit for the regression line. Relying on the initial estimate for uninfected subsamples would have lead us to conclude incorrectly that an entire host was uninfected. We avoided this potential pitfall by increasing the sampling effort until parasites were found or until the entire uninfected host had been pro-

cessed. Other subsampling techniques may not encounter a problem with uninfected subsamples.

Infection dynamics will dictate what type of predictive model can be used when designing a subsampling method. With bilaterally symmetrical infection sites, a simple ratio model may be sufficient for estimating mean intensity (e.g., double the subsample count to obtain total intensity). Asymmetrical or multiple infection sites will require another method (e.g., the regression model delineated in the present paper). Parasites that occupy more than 1 organ present an additional complexity in which several subsampling strategies may be available. With a linear regression model, these potential strategies can be compared to find the one that maximizes the ratio $R^2(c_t + c_s)^2/(4c_tc_s)$. Furthermore, the efficiency of a proposed model should be evaluated before it is employed to ensure that $R^2 > 4c_tc_s/(c_t + c_s)^2$. Although we were able to fit a linear regression model to sample up to 3 parasite species in a single host with a unified approach, initially the most efficient subsampling routine differed among parasites for each fish host. This indicates that such time saving methods are species-specific at the host and parasite levels. We suggest these unique infection patterns demonstrate that tissue generalists do not treat all organs alike and that their distribution within a particular fish is (host) species-dependent.

Our linear regression model was developed to increase the efficiency of sampling large numbers of hosts over the extent of a long-term study. The targeted host–parasite systems are found at multiple sites along a large geographic range. A subsampling method generated using data from one site, however, should not be employed at a different site without verification that the slopes and intercepts of the regressions do not vary among sites. This requires a complete examination of an initial sample of hosts to verify that the relationship between total intensity (I) and the intensity observed in the subsamples (S) is the same for both initial and subsequent samples ($\hat{I}_2 \approx b_1S_2 + a_1$). Additionally, the model will need to be recalibrated because of factors such as seasonal variation or the detection of cryptic parasite species (this method assumes that all species are known and identified).

In conclusion, subsampling can aid in the examination of a large number of hosts harboring high-intensity infections. Although choosing an effective subsampling routine becomes a complex process when considering trade-offs in accuracy and effort reduction, the optimal strategy may produce long-term rewards in the form of time saved. Researchers can devote their additional free time to processing larger sample sizes or other endeavors (at their discretion).

ACKNOWLEDGMENTS

LITERATURE CITED

BUSH, A. O., K. D. LAFFERTY, J. M. LOTZ, AND A. W. SHOSTAK. 1997. Parasitology meets ecology on its own terms: Margolis et al. revisited. Journal of Parasitology **83:** 575–583.

COCHRAN, W. G. 1977. Sampling techniques. Wiley, New York, New York, 428 p.

GRACZYK, T. 1991. Cases of bilateral asymmetry of *Diplostomum pseudospathecum* Niewiadomska, 1984 metacercariae infections (Trematoda, Diplostomidae) in the eye lens of fish. Acta Parasitologica Polonica **36:** 131–134.

MARCOGLIESE, D. J., P. DUMONT, A. D. GENDRON, Y. MAILHOT, E. BERGERON, AND J. D. MCLAUGHLIN. 2001. Spatial and temporal variation in abundance of *Diplostomum* spp. in walleye (*Stizostedion vitreum*) and white suckers (*Catostomus commersoni*) from the St. Lawrence River. Canadian Journal of Zoology-Revue Canadienne De Zoologie **79:** 355–369.

MARTIN, W. E. 1950a. *Parasitictodora hancocki* N-Gen., N-Sp. (Trematoda, Heterophyidae), with observations on its life cycle. Journal of Parasitology **36:** 360–370.

———. 1950b. *Phocitremoides ovale* N. Gen, N. Sp (Trematoda, Opisthorchiidae), with observations on its life cycle. Journal of Parasitology **36:** 552–558.

SOKAL, R. R., AND J. R. ROHLF. 1981. Biometry. W.H. Freeman and Company, New York, New York, 859 p.

THIEMANN, G. W., AND R. J. WASSERSUG. 2000. Biased distribution of trematode metacercariae in the nephric system of *Rana* tadpoles. Journal of Zoology **252:** 534–538.

YAMAGUTI, S. 1971. Synopsis of digenetic trematodes of vertebrates. Keigaku Publishing Co., Tokyo, Japan, 1074 p.