ACCURACY IN MOOSE MANDIBLE SIZE VERSUS TOOTH WEAR ASSESSMENTS

Brian McLaren and Richard Curran

Government of Newfoundland & Labrador, Department of Forest Resources & Agrifoods, Inland Fish & Wildlife Division, P. O. Box 8700, St. John's, NF, Canada A1B 4J6

ABSTRACT: Accuracy checks of tooth wear assessments and size measurements were undertaken for mandibles submitted by moose hunters in Newfoundland. Tooth wear class, a subjective assessment often used in age interpretation, was inconsistent in repeat measurements by the same technician in 23% of cases, and, in comparison between 2 technicians, in 53% of cases. There was confusion particularly between Classes II and III, and consistent bias in the interpretation of higher classes. Size measurement was much more consistent, with 1-2% relative error in repeat measurements of a sample size of $n \geq 77$. Smaller sample sizes may result in larger relative measurement error because individual repeat measurements differed by as much as 17 mm.

ALCES VOL. 37 (1): 13-17 (2001)

Key words: age assessment, jaw condition, measurement error, moose, Newfoundland, tooth wear

Moose (*Alces alces*) populations are best managed with accurate sex and age information (Bubenik et al. 1975, Ferguson 1993). When moose mandibles are submitted to research agencies by hunters, assessment of the jaws and teeth can be used to determine age structure of the harvested populations. While this strategy allows monitoring at a low cost, it depends on accurate age assessments. Assessments by most wildlife agencies include cementum age from incisors (Haagenrud 1978), tooth eruption, wear or jaw condition (Passmore et al. 1955), and mandible size (Saether 1983). When an agency is responsible for a large moose harvest, or a harvest over a large geographic area, several wildlife officers or technicians may be involved in processing mandibles. This paper addresses the accuracy of such assessments.

METHODS

In Newfoundland, moose mandible submissions are required as part of a Big Game Hunter participation and recognition program. Mandibles are stored frozen until processing, and several thousand assessments are made each hunting season by staff in a central laboratory facility in St. John's, Newfoundland, as well as in some regional centres in the Province. Due to limited storage space, mandible ages and corresponding licence numbers must be recorded as soon as possible after they are received. To minimise observer error and standardise subjective wear assessments as much as possible, an attempt is made each year to have as few technicians as possible (2 or 3) process mandibles, and to have the same technician assess tooth wear each year. Sometimes, accuracy in assessments may be sacrificed for processing efficiency.

Age is easily recognised from mandibles taken from calves and yearlings, while mandibles from animals ≥ 2 years old are limited to a tooth wear assessment, using Classes II through IX of the 10-class system developed by Passmore et al. (1955). This method is based on exposure of light

and dark dentine patterns on premolar and molar crowns, recession of gums around molars, and a subjective assessment of molar concavity, or "scooping." To assist wear assessments, photographs from Passmore et al. (1955) with accompanying descriptive text are available for frequent consultation. Resulting age and wear assessments are recorded with hunter licence number. All jaws are then measured for total lower mandible length. Total length is that measurement from the buccal side of the most proximal part of the jaw bone to the most distal part of the jaw bone at the furthest extent of the gum line (Fig. 1). Recession and presence of gum tissue is variable depending on cleaning and storage of the bones, and estimates of the gum line are approximate. Additional mandible measurements taken from moose ≥ 2 years old are molar row length, from the proximal gum line at the third molar (M3) to the distal gum line at the first premolar (P2), and diastema length, from the distal gum line at the first premolar (P2) to the proximal gum line at the fourth incisor (I4) (Fig. 1). All measurements are recorded to the nearest millimetre with a steel tape, and wherever possible, measurements are made on the left lower mandible (if the left is broken, the right is substituted). Measurements are true lengths, placing the tape straight and parallel to the jaw as much as possible. Two incisors (I1) are extracted from animals ≥ 2 years old for cementum ageing. This paper does not explore relationships between wear class and cementum age, because the subject has been handled in more detail elsewhere (Addison and Timmermann 1974, Hindelang and Peterson 1993).

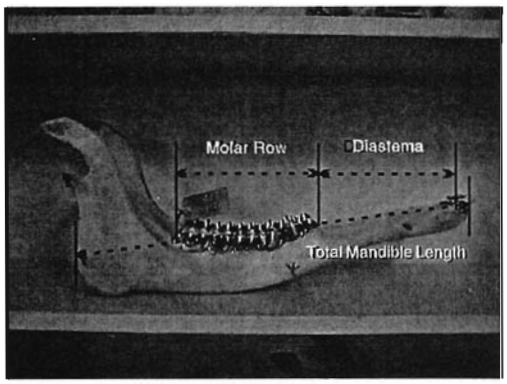

Fig. 1. Endpoints (—) and dimensions (- - - -) measured on mandibles submitted by moose hunters in Newfoundland. See Methods for descriptions of these measurements.

Ninety-six mandibles were compared, using a blind-check routine, including 4 calf and 10 yearling mandibles. After initial assessments by one technician, all mandibles were replaced in storage bins and reassessed by a second technician with less experience. The second technician used a different steel measuring tape. After the second assessment, all mandibles were remeasured by the first technician. The 3 sets of observations for each mandible were then matched using licence number. Wear assessments were compared by cross-tabulation. Size measurements were compared with the assumption of a normal error distribution.

RESULTS

Cross-tabulation of wear assessments indicated a 23% difference in reassignment of tooth wear classes by the same technician, and a 53% difference between the assignments by 2 technicians (Table 1). The majority of difference between observers occurred in Classes II and III, and when these classes (according to the first assessment) were removed, the results became 20% and 31% error respectively. Comparing the first and second assessments by the same observer, about the same number of mandibles were assigned a higher wear class (n =13) as a lower wear class (n = 8) the second time. Comparing the two independent assessments, tooth wear in the initial assessment in Classes III and higher was more often assigned a lower class by the second observer, in 29 of 50 cases,

Table 1. Comparison of initial wear class assignments for moose jaws (lower mandibles) from eastern Newfoundland with (a) repeated assignments by the same technician, and (b) assignment by a second technician.

Initial Wear Class	(a) Repeated assignment								(b) Second assignment					
Assigned	I	II	III	IV	V	VI	VII	VIII	I	II	III	IV	V	VI
I	12								10	3				
II	1	23	5							13	15	1		
III		3	13	3	1					10	8	1		
IV		1	1	12						2	8	5		
V					5	2						2	5	
VI					1	2	2				1	2		2
VII						1	1					1	1	1
VIII								1				1		

while a higher wear class was assigned only once in 50 cases (i.e., observer bias; total difference between observers for jaws in Class III or higher, 60%). The 4 calf mandibles were consistently identified, but yearling assessments varied, with 1 of 13 falling into Class II in reassessment by the same observer, and 3 of 13 falling into Class II in assessment by different observers.

The results of the measurement error assessment showed an encouraging level of precision and accuracy in both observers (Table 2). Comparing the blind check on the first observer, mean error was < 2 mm for all 3 measurements. The 95% confidence interval for the absolute error in this assessment was approximately 1-2 mm, and relative error was ≤ 1%. Moreover, from the sample of mandibles used in this study, 95% of the remeasurements for this observer fell well within ± 1 mm. Comparing observers, the error rate reflected a consistent negative bias of 2-3 mm, and the absolute measurement error was approximately 3-4 mm (95% C.I.) or 1-2% relative error.

DISCUSSION

The results of this study describe the subjectivity of the wear class index. Misclassification even of yearling moose is not unexpected, based on age comparisons in Hindelang and Peterson (1993), who showed 30 moose yearlings in a sample of 76 aged by cementum annuli being determined as Class II or higher. Classes II and III, which accounted for the majority of difference between observers in this study, illustrate the difficulty in wear assessment. These 2 classes are distinguished only by a slight difference in wear on the third molar, a difference in breadth of exposure of darker dentine on the first molar, and a subjective assessment of wear on the premolars (Passmore et al. 1955). Because tooth wear is progressive, subjective assessments will differ considerably between observers. When technicians are most experienced in assessing samples that contain few mandibles of relatively old animals, observer inexperience with excessive wear explains the bias found between technicians in classifying higher than Class II (cf. Hamlin et al. 2000). It is encouraging to note that reas-

Table 2. Measurement error, mm, in (a) repeated observations by the same technician, and (b) repeated measurement by a second technician, for mandible size in adult moose from eastern Newfoundland.

	(a) Same technician			(b) Second technician		
	Total	Molar Row	Diastema	Total	Molar Row	Diastema
Number of cases	80	78	77	83	80	79
First mean	432	165	154	432	165	155
Second mean	432	165	155	431	162	152
Mean of paired differences	-0.238	-0.026	0.169	-2.346	-2.722	-2.671
95% C.I. for paired differences	-0.851, 0.376	-0.503, 0.452	-0.351, 0.688	-3.195, -1.496	-3.461, -1.982	-3.313, -2.029
Range of absolute measurement error	0-9	0-11	0-9	0-17	0-14	0-14
Mean absolute measurement error	1.912	1.333	1.597	3.383	3.557	3.101
95% C.I. for measurement error	1.464, 2.361	0.960, 1.706	1.220, 1.974	2.723, 4.043	3.019, 4.094	2.563, 3.640
Relative error (%)	0.4	0.8	1	0.8	2.2	2

sessment by the same observer only rarely involves a change in classification of more than one wear class (Table 1), but when this difference involves yearlings and 2-year-old moose, population reconstruction is seriously affected, because these are such a large component of most managed moose populations. Also, comparison of the results between observers offers little justification for comparing moose population age structures based on tooth wear classes if more than one technician is involved in a study. Similarly, patterns of tooth wear cannot be compared between populations if different technicians are involved in the assessment. Finally, tooth wear assessment is not a precise tool, and should not be used as a surrogate for ageing moose by counting tooth cementum annuli (Hamlin et al. 2000).

Measurement of mandible size is, in contrast, a fairly reliable procedure. Between years, technicians involved in mandible processing may change and still allow an acceptable accuracy for the comparison of populations or population cross-sections. The consistent (negative) bias in this study in the results of 2 observers is probably related to differences in the measuring tape, or its use (e.g., parallax error). It is important that a large enough sample be used in making mandible size comparisons between populations, because, as this study shows, single measurements can involve errors of up to 17 mm; much higher than the mean relative error of 1-2% (Table 2). These

large errors are probably associated with incomplete cleaning of tissue from the jaw or with speed in processing mandibles. This study also indicates that reported differences in mandible size of < 5 mm are likely associated with measurement error. It is unfortunate that mandible size cannot easily be translated into age classes for moose.

ACKNOWLEDGEMENTS

Barry Adams of the Newfoundland and Labrador Department of Forest Resources and Agrifoods, Inland Fish & Wildlife Division, has coordinated the laboratory technicians each year for mandible measurement since 1994, and offered extra assistance for this study. As in each hunting season, this Department is very grateful for this assistance, and for the participation of so many moose hunters and Conservation Officers in the submission, collection, and assessments of mandibles.

REFERENCES

ADDISON, R.B., and H.R. Timmermann. 1974. Some practical problems in the analysis of the population dynamics of a moose herd. Proceedings of the North American Moose Conference and Workshop 10:76-106.

BUBENIK, A. B., H. R. TIMMERMANN, and B. SAUNDERS. 1975. Simulation of population structure and size in moose on behalf of age structure of harvested animals. Proceedings of the North American Moose Conference and Workshop 11:391-463.

FERGUSON, S. H. 1993. Use of cohort analysis to estimate abundance, recruitment and survivorship for Newfoundland moose. Alces 29:99-114.

HAAGENRUD, H. 1978. Layers in secondary dentine of incisors as age criteria in moose. Journal of Mammalogy 59:857-858.

HAMLIN, K. L., D. F. PAC, C. A. SIME, R. M. DESIMONE, and G. L. DUSEK. 2000. Evaluating the accuracy of ages obtained by two methods for Montana ungulates. Journal of Wildlife Management 64:441-449.

HINDELANG, M., and R. O. PETERSON. 1993. Relationship of mandibular tooth wear to gender, age, and periodontal disease of Isle Royale moose. Alces 29:63-73.

PASSMORE, R. C., R. L. PETERSON, and A. T. CRINGAN. 1955. A study of mandibular tooth wear as an index to age of moose. Pages 223-238 *in* R. L. Peterson. North American Moose. University of Toronto Press, Toronto, Ontario, Canada.

SAETHER, B. E. 1983. Relationship between mandible length and carcass weight of moose in Norway. Journal of Wildlife Management 47:1226-1229.