REPEATABILITY OF INCISOR CEMENTUM AGE DETERMINATIONS FOR MOOSE

H. G. Cumming, School of Forestry,
Lakehead University, Thunder Bay, Ont.

R. Evans, The Proctor & Redfern Group,
Toronto, Ont.

Abstract: The incisor cementum method for determining moose ages failed to yield consistent results in three separate tests. Only 9 of 50, 11 of 150 and 3 of 50 incisors were assigned identical ages by the 3 to 4 observers in each test. The differences exceeded those of a single observer comparing the incisor cementum method with the wear class method. Mean ages, percentages of yearlings in the harvest and deviations from regression of mortality curves all varied significantly among observers, suggesting the possibility of serious errors in management decisions. Special guidelines failed to improve consistency but two biologists taught by the same person and consulting on interpretation obtained close agreement. We recommend annual training sessions for standardizing interpretation within each jurisdiction.

The discovery of cementum rings in moose (Alces alces) incisors by Sergeant and Pimlott (1959) introduced a new era in moose aging. Many advantages were immediately evident. No longer was it necessary to collect and store huge piles of lower jaws in order to use the wear class method developed by Passmore, Peterson and Cringan (1955). No more training sessions were required to standardize use of rather nebulous criteria (Cumming 1974); technicians only needed to count the

rings to determine the ages of moose. Perhaps most importantly, the new method was held to be more accurate, Simkin (1968) comparing three methods of aging moose used the terms "cementum ring count" and "actual age" synonomously. Addison and Timmerman (1974) looked for ways of converting wear class statistics into forms comparable with incisor cementum data but found it was not practical beyond 2 1/2 years due to variations which they ascribed to the wear class method. Dzieciolowski (1976) made a similar comparison of aging techniques for European moose (Alces alces alces) and produced a conversion equation, again assuming that the incisor method was correct. Sergeant and Pimlott (1959) had one tooth of known age. No one tried to find how accurate the method really was until Gasaway, Harkness and Rausch (1978) accumulated 68 moose of known ages. In addition to their main study they included age determinations by two of the authors to illustrate the variability of age estimates between observers. No other studies of repeatability of the incisor aging method have been reported.

Repeatability was the subject of this study. The questions we attempted to answer were the following: (1) Is the cementum aging method really reliable? (2) What differences may be found among observers? (3) What features of teeth make age determination difficult? (4) What rules will ensure that everyone interprets what he sees as much as possible in the same way? (5) How much improvement can be expected?

METHODS

The first test of repeatability was carried out by the senior author in 1960 by (a) sorting 50 moose jaws into piles of jaws with similar wear classes, then assigning ages according to Passmore et al. 1955 and (b) extracting the primary incisors (I_1), grinding away half the roots with an emery stone, then examining the resulting longitudinal sections under a microscope using strongly directional, oblique, reflected light. After assigning ages to the entire set of incisors on two separate occasions several days apart using the criteria described by Sergeant and Pimlott (1959), he then sent the incisors to two other biologists for age determinations with the same paper as a guide.

The study was resumed in 1973 when the junior author's assistance was provided by the Ontario Ministry of Natural Resources (OMNR). For this second test, samples of moose jaws collected from hunters by OMNR staff in Swastika and Kapuskasing districts of Ontario were first assigned ages as before using the wear class method. Then both authors together assigned ages using the incisor cementum technique as already described, except that a jewellers saw was used to split the root of each tooth longitudinally. Following initial aging, the best incisor sections from Swastika were divided into 3 sub-samples of 50 each according to the ease with which ages could be assigned. These became the incisors for the second test of repeatability. They were numbered, and sent in turn to three Ontario districts where ages were assigned by district staffs using normal procedures. We intended to change numbers on the incisors between each district determination to prevent any sequential bias but through misunderstanding this was not done.

A third test was carried out in 1977 after analysing the accumulated data and preparing guidelines aimed at standardizing interpretation by observers. In this test the incisors from the

Swastika sample that were considered neither especially easy nor especially difficult for age determination were assigned ages twice by the senior author, once each by two inexperienced volunteers and once by an experienced OMNR technician using the proposed guidelines.

The incisors from Kapuskasing were retained for measuring cementum widths using a measuring eye piece in a dissecting microscope at 50 X magnification. Some were then sent to the wildlife laboratory of the OMNR Research Branch at Maple, Ontario, for decalcification, staining and preparation as microscope slides for use with transmitted light.

RESULTS

The 1960 Test

The first test of repeatability (1960) showed that only 8 of the 50 incisors were assigned identical ages in all four determinations (two by the senior author and one each by the other two biologists). Although 23 of the assigned ages differed by only 1 year, 15 differed by 2 years, 4 by 3 years and 1 by 5 years. Examination of these results in pairs (Table 1) showed that correlation R-values were universally high (0.930 or greater), but less than half of the assigned ages were identical for any pair of observations. Although most disagreements were by only one year, there were also a substantial number of greater disagreements (Table 1). The second determination by the senior author showed no consistent bias toward over or under estimation compared with the first one but the other comparisons all showed highly significant differences among observers. Most of the ages assigned by one biologist were higher than those assigned by the senior author whereas most by the other

Table 1. Results of 1960 repeatability test for the incisor cementum method of determining moose ages (N=50)

	Paired Comparisons					
	A1/A2 *	A1/B	A1/C	A2/B	A2/C	B/C
Frequency distribution of age differences (yrs)						
0	24	24	24	20	21	17
1	18	20	19	26	18	22
2	7	5	5	2	9	8
3	1			1	1	2
4			1		1	1
5		1		1		
Sum of + and - signs	0	-16	+20	-18	+13	+33
Correlation R-values	.960	.952	.952	.946	.930	.957
Mean difference (yrs)	0	-.500	.480	-.520	.420	.960
t-value of difference	0	3.416	3.711	3.256	2.481	7.339
Probability of difference departing from o	o	<0.005	<0.001	<0.005	< 0.025	<0.001

* A indicates the senior author, with subscript indicating two different determinations in the 1960 test.
B and C refer to the other biologists who took part in the test.

biologist were lower. These two biologists differed very significantly (t = 7.339, P<0.001) with all differences in the same direction (Table 1).

The 1973 Test

In 1973 after moose biologists had had 14 years' experience with the incisor cementum method, some improvement was evident. The greatest difference among assigned ages was only 3 years rather than 5 (Table 2) and the correlation R-values were even higher. But only 11 of the 150 teeth examined in the second test were assigned the same age in all 4 determinations. Another 86 were assigned ages only 1 year apart; 45 were 2 years different; and 8 were 3 years different. Agreement between pairs of participants varied greatly. Most showed strong, consistent biases toward older or younger ages, but one pair (D and F) showed practically no bias (Table 2). Most differences between observers were highly significant (P<0.001) but D and F showed small probability of being different (P>0.500). The two agreeing biologists assigned 130 identical ages to the 150 teeth. Rather less agreement was documented between the authors and biologist E (69 identical determinations out of 150). But the greatest divergences occurred between the authors' determinations and those of the district biologists who showed close agreement (D and F).

These differences raised the question of which interpretation of the cementum observations to accept. The close agreement between agers D and F suggested that they might be right. On the other hand, comparisons between author-assigned ages and ages determined by the wear class method for the same jaws, including the easily

Table 2. Results of 1973 test of repeatability in determining moose ages by the incisor cementum method.

(a) Incisors considered easy for determining ages (N=50)

	Paired Comparisons					
	A4/D	A4/E	A4/F	D/E	D/F	E/F
Frequency distribution of age differences (yrs).						
0	1	20	1	23	45	23
1	39	28	39	25	5	24
2	10	2	9	2		3
3			1			
Sum of + and - signs	+49	+28	+49	-27	+ 1	+27
Correlation R-values	0.989	0.984	0.988	0.981	0.994	0.979
Mean difference &yrs)	1.18	0.60	1.20	-.58	.02	.60
t-value of difference	11.306	6.980	17.145	7.137	0.444	7.004
Probability of difference departing from 0	<0.001	<0.001	<0.001	<0.001	>0.500	<0.001

*A4 refers to the authors. Other letters refer to district personnel taking part in the test.

Table 2 continued.

(b) Incisors considered difficult for determining ages (N=50)

	Paired Comparisons					
	A4/D	A4/E	A4/F	D/E	D/F	E/F
Frequency distribution of age differences (yrs)						
0	2	25	3	9	43	9
1	31	25	29	29	7	29
2	14		15	11		10
3	3		3	1		2
Sum of + and - signs	+48	+15	+47	-39	+ 1	+39
Correlation R-values	0.910	0.976	0.968	0.965	0.991	0.962
Mean difference (yrs)	1.38	0.30	1.36	-1.04	0.02	1.06
t-value of difference	14.630	3.282	13.881	9.743	0.375	9.452
Probability of difference departing from 0	<0.001	<0.005	<0.001	<0.001	>0.500	<0.001

Table 2 coninued.

(c) Incisors considered neither especially easy nor difficult for determining ages (N=50).

	Paired Comparisons					
	A4/D	A4/E	A4/F	D/E	D/F	E/F
Frequency distribution of age differences						
0	9	24	12	22	42	22
1	30	23	26	22	7	24
2	10	3	11	5	1	4
3	1		1	1		
Sum of + and - signs	+33	+16	+32	-20	0	22
Correlation R-values	0.943	0.961	0.958	0.950	0.984	0.963
Mean difference (yrs)	0.90	0.38	0.90	-0.50	-0.02	0.48
t-value of difference	7.183	3.568	7.179	3.989	0.299	4.636
Probability of difference departing from 0	<0.001	<0.001	<0.001	<0.001	>0.500	<0.001

identified Wear Class I (1 1/2 years) tended to support the authors' age determinations. To obtain more information incisor ages assigned to teeth from Wear Class I and II jaws were examined separately. Only 1 of 43 Wear Class I jaws had been assigned the age of 2 1/2 years rather than 1 1/2 years by the incisor cementum method. Furthermore, only 2 of 25 Wear Class II jaws were aged as 1 1/2 years rather than 2 1/2 years. Therefore, errors between 1 1/2 and 2 1/2 years by the authors were small.

Another approach was to look at cementum widths obtained from the Kapuskasing sample. A bar graph showed an initial high frequency of measurements just before 0.4 mm in width followed by relatively few measurements until 0.6 mm suggesting that yearling moose might be those under 0.4 mm and 2 1/2 year-old-moose those at 0.6 mm. A re-check of 46 teeth from the Kapuskasing sample with cementum widths of 0.4 mm or less showed that all but 2 had been assigned the age 1 1/2 years by the authors, again confirming the accuracy of the authors' age determinations for 1 1/2 and 2 1/2 years. Examination of the ages assigned by the other biologists showed that only 7 of the 34 incisors from the 1973 Swastika sample that the authors believed to be 1 1/2 years old, were designated as such by biologists D and F, the others being called 2 1/2 years old. This was in contrast with biologist E who agreed with the authors in calling 32 of the 34 incisors 1 1/2 years old. Therefore, it appears that although biologists D and F showed close agreement, their age determinations were in fact inaccurate. In a trial to improve agreement with the ages assigned by the authors, one year was subtracted from each of the age determinations by biologists D and F. The differences then decreased as follows:

	Years of disagreement 0	1	2	3
Differences between ages assigned to the 1973 Swastika sample by D and F compared with those of the authors	14	97	35	4
Differences when one year was added to each of the ages assigned by D and F	98	42	10	

Obviously, there was a consistent tendency by the two biologists to over-age the incisors by at least one year.

To find why biologists D and F achieved close agreement while biologist E and the authors differed substantially from them, we wrote all three biologists asking how they carried out the aging and how they learned the techniques. Agers B and D aged the sample independently but at the same time and place using the same equipment. After completing their determinations they discussed interpretation of ages on which they disagreed and changed a few but continued to disagree on most. Both had learned the technique from the same person. Biologist E had learned the method from another biologist and had not conferred with D and F. All three had Sergeant and Pimlott's (1959) original article but did not necessarily follow it closely.

The 1977 Test

Results of the test using guidelines (Appendix 1) appeared even worse than previous comparisons (Table 3). Only 3 of 50 teeth were assigned identical ages all 5 times (twice by the senior author and once each by two inexperienced and one experienced ager). For 15 incisors the difference in assigned ages amounted to only 1 year; for another 19 it was 2 years; for 6, 3 years; for 3, 4 years; for 3, 5 years and for 1, 6 years. The paired analysis showed lower correlation R-values

Table 3. Results of 1977 test of repeatability in determining moose ages by the incisor cementum method using proposed guidelines (N=50)

	A4*/A5	A4/G	A4/H	A4/I*	A5/G	A5/H	A5/I	G/H	G/I	H/I
Frequency distribution of age differences (yrs)										
0	24	22	19	19	18	16	24	15	14	15
1	21	16	20	21	18	25	17	18	22	25
2	4	7	5	6	9	6	6	6	7	3
3	1	1	4	3			1	4	4	1
4		2	1	**	3	1		4	2	1
5		2	1		1	2	1	2	**	3
6					1		**	1		1 **
Sum of + and - signs	2	16	-6	12	14	-6	11	-11	-5	14
Correlation R-values	.919	.840	.785	.910	.811	.810	.907	.772	.845	.815
Mean difference (yrs)	.04	-.74	.42	-.41	-.78	.38	-.43	1.16	.37	-.80
t-value of difference	0.292	3.531	1.998	2.489	3.353	1.812	2.566	4.529	1.701	3.256
Probability of difference departing from 0	>0.500	<0.001	<0.100	<0.025	<0.005	<0.100	<0.025	<0.001	<0.100	<0.005

* A_4,A_5 refer to the senior author, G, H to inexperienced agers, I to an experienced ager.

** One incisor not aged by I

(Table 3) than in the other tests and consistent biases between observers were evident in all but 4 comparisons, one of those being the senior author's repeated determinations, two being comparisons between the senior author and one inexperienced observer and one being a comparison between an inexperienced observer and the experienced observer. Differences between the repeated determinations by the senior author were almost identical to those in the first test 18 years earlier. In fact, there seemed no general improvement in determining moose ages during this time period since the mean number of identical determinations was 21.6 for the 1960 test, 18.8, 15.2 and 21.8 for the sub-samples of the 1973 test and 18.0 for the 1977 test.

The question remains whether these differences in age determination have any practical significance. Aging errors which did not affect parameters commonly used in moose management would be of little consequence. Management uses of moose age data include (1) calculation of mean age (Karns et al. 1974), (2) percentages of yearlings in the harvest (Ritcey 1974, Cumming 1974) and (3) survival tables or mortality curves (Karnes et al. 1974, Cumming 1975). Differences among mean ages for the 150 incisors in the 1973 test (4.29, 5.42, 4.71 and 5.42) proved to be highly significant (F=6.33, P<.01) as were percentages of yearlings (x^2=39.15 P<.01). The slopes of the two most different mortality curves after logarithmic transformation were not significantly different (F=0.6115 P<.05) but one deviated from regression significantly more than the other (F=10.25 P<.01). Thus all management parameters were altered significantly by differences in age determination among observers.

Table 4. Comparisons between the incisor cementum method of determining moose ages and the wear class method,* using the jaws from both the first and second test (1960,1973).

	Paired Comparisons		
	1960		1973
	A1/WC	A2/WC	A3/WC
	(N=50)	(N=50)	(N=143)
Frequency distribution of age differences (yrs)			
0	28	29	123
1	17	14	15
2	5	6	3
3			2
4		1	
Sum of + and - signs	+12	+11	- 7
Mean difference (yrs)	0.30	0.32	-0.05
+-value of difference	2.605	2.266	1.044
Probability of difference departing from 0	<0.025	<0.050	<0.400

*Ages assigned each wear class followed Passmore et al (1955) as follows:

WCI - 1 1/2 yrs, WCII - 2 1/2 yrs, WCIII - 3 1/2 yrs.

WCIV - 4 1/2 yrs, WCV - 5 1/2-6 1/2 yrs, WCVI - 6 1/2-8 1/2 yrs, WCVII - 8 1/2-10 1/2 yrs, WCVIII - 10 1/2-15 1/2 yrs, WCIX - 14 1/2 + yrs.

Comparisons with the wear class method (1960 and 1973 results)

Over half the incisors in the 1960 sample fell within the ranges of ages suggested by Passmore et al. (1955) for their described wear classes (Table 4). The greatest difference was 4 years. A significant bias was evident between the two methods with the wear class method tending to underestimate the ages of older moose. However, in the 1973 test 123 incisors out of 150 were assigned ages within the ranges assigned to the corresponding jaws by the wear class method and there was no significant bias. Therefore, differences between the incisor cementum method and the wear class method by a single observer were fewer than differences between observers using the incisor cementum method.

The wear class method has always been assumed less accurate for older moose (Passmore et al. 1955, Simkin 1968) but repeatability does not decrease with age when using the incisor cementum method. In the 1960 test the number of identical determinations was about as frequent among older age classes as among younger ones (Table 5). In the 1973 test, agreements were too few for similar statistics but disagreements showed only a slight increase with age in both the mean number of differences per tooth and the mean difference per tooth in years (Table 6). Efforts to improve the incisor cementum method by using decalcified, stained microscope sections were unsuccessful. In fact, cementum layers were more difficult to count.

Table 5. Analysis of the effect of moose ages on agreements among observers in the first test (1960)

	Ages in years								
First age assigned by senior author	2½	3½	4½	5½	5½	7½	8½	9½	10½
Number of incisors in each age class	17	6	7	5	3	2	2	3	5
Mean number of identical determinations per tooth out of possible 3	2.39	1.43	1.57	2.20	1.33	2.00	2.40	3.00	1.29

Table 6. Analysis of the effect of moose ages on differences among observers in the second test (1973)

Age agreed on by both authors	1½	2½	3½	4½	5½	6½	7½	8½	9½	13½
Number of incisors in each age class	23	20	15	7	9	8	6	7	4	1
Mean number of different determinations per tooth out of a possible 3	1.57	2.80	2.53	2.43	2.67	2.75	2.67	2.71	2.50	1.0
Mean difference in assigned age in years per incisor	2.01	3.30	3.07	2.86	3.33	3.38	4.00	3.71	8.25	4.0

DISCUSSION

The results indicate that early enthusiasm for the incisor sectioning method of aging moose was a little exaggerated. Advantages in collecting and storing incisors remain, but ages determined from incisors cannot be considered exact. Some of the discrepancies previously discovered between the incisor method and other aging methods were undoubtedly due to errors in the incisor method; they were not all one-sided as then assumed.

More is involved in aging incisors than merely being able to count. The greater variation among inexperienced agers suggests that practice helps; but the consistent differences among all but two people taking part in these tests show that agreement on interpretation is necessary. The failure of both Sergeant and Pimlott's (1959) original paper and our guidelines to produce consistent interpretation of what was seen suggests that written instructions alone are not enough. Best agreement occurred when agers examined the incisors together and consulted on their findings. Gasaway (pers. comm.) reported similar close agreement among people whom he had trained himself. Therefore, it seems that training sessions were discontinued prematurely. Probably each State or Province will have to re-introduce training sessions in which inexperienced agers gain vital experience and experienced agers standardize interpretations.

It is possible that these incisors were exceptionally difficult to age, but attempts to improve agreement by eliminating difficult incisors proved unsuccessful. Therefore, the significant differences among management parameters suggest that serious errors in judgement

could result from aging errors during standard management applications. A decrease of more than one year in the mean age coupled with a significant drop in the percentage of yearlings would appear quite alarming to someone responsible for adjusting moose harvest intensity and might result in closed seasons even though the changes were entirely due to aging errors. When one considers along with these results the 56% difference from known age for animals of 2 years and older found by Gasaway et al. (1978), one wonders whether the incisor aging method is worth pursuing at all. Perhaps one advantage of the wear class aging method was that everyone knew it was inaccurate, whereas the incisor sectioning method gives the appearance of accuracy. We believe the incisor method is still worth using but more attention needs to be given to standardization and better guidelines need to be developed from known-aged animals if it is to become a fully reliable technique.

ACKNOWLEDGEMENTS

We would like to acknowledge the help of those who cooperated in these aging tests. They are too numerous to mention individually but without them there would have been no study. We would also like to acknowledge the support of the Ministry of Natural Resources with whom we were employed during the major part of this study. We are grateful to William Gasaway for valuable comments on statistical treatments and the text.

REFERENCES

Addison, R.B. and H.R. Timmermann. 1974. Some practical problems in the analysis of the population dynamics of a moose herd. Proc. 10th N.A. Moose Workshop, Duluth, Minn. pp. 76-106.

Cumming, H.G. 1974. Annual yield, sex and age of moose in Ontario as indices to the effects of hunting. Naturalists Can. 101: 539-558.

Cumming, H.G. 1975. Analysis of long-term moose hunting statistics from Ontario. Proc. 11th N.A. Moose Conf. and Workshop. pp. 218-244.

Dzieciolowski, Ryszard. 1976. Age appraisal in moose. Acta Tehriologica 21, 23: 307-310.

Gasaway, William C., David B. Harkness, Robert A. Rausch. 1978. Accuracy of moose age determinations from incisor cementum layers.

Karns, P.D., H. Haswell, F.F. Gilbert, and A.E. Patton. 1974. Moose management in the coniferous-deciduous ecotone of North America. Naturalists Can. 101: 643-656.

Passmore, R.C., R.L. Peterson and A.T. Cringan. In Peterson, Randolph L. 1955. North American Moose. U. Toronto Press. Toronto.

Ritcey, R.W. 1974. Moose harvesting programs in Canada. Naturalist Can. 101: 631-642.

Sergeant, D.E. and D.H. Pimlott. 1959. Age determination in moose from sectioned incisor teeth. J. Wildl. Manage. 23(3): 315-321.

Simkin, D.W. 1968. A comparison of three methods used to age moose. Proc. 5th N.A. Moose Workshop, Kenai, Alaska. pp. 46-59.

APPENDIX I

GUIDELINES FOR INTERPRETING OBSERVATIONS OF INCISOR CEMENTUM WHEN DETERMINING AGES OF MOOSE

To produce age determination guidelines we drew up a set of rules defining (1) exactly where and how to begin counting cementum layers and (2) what to do when cementum layers split. The first point seemed important because wide disagreement occurred even when deciding between 1 1/2 and 2 1/2 year-old-moose. The second problem came to light when we examined 33 incisors for which ages using the incisor cementum method differed substantially from wear-class assigned ages. In 13 cases cementum bands seemed split along their length making more than one age determination possible. In 7 of those cases the splits were near the middle of the incisor and in 6 they were near an end. Langenau (1972) observed similar splitting problems in white-tailed deer. We found that agreement with the wear class method improved in 10 of the 13 cases when splits near the middle of the incisor were disregarded but those near an end counted. Therefore, the guidelines recommended that this be done. The following "rules" were provided to each person taking part in the last test:

REFERENCE

Langenau, E.E. 1972. Applications and limitations of aging white-tailed deer annuli in the cementum of the first lower incisor. Research and Development Report No. 272, Michigan Dept. Nat. Res.

Appendix I

RULES FOR AGING MOOSE INCISORS

H. G. Cumming, January, 1978.

1. To examine under a binocular microscope set tooth in plasticene or similar substance.

2. Look along the cementum layers on both sides of the tooth root using a low power light with strong oblique light. Locate the side of the incisor with the greatest cementum width and find the widest point on that side with clear cementum banding. You will want to record the maximum number of cementum layers as it will always be possible to find fewer layers elsewhere on the tooth.

3. Locate the cemento-dentine juncture, usually indicated by a prominent white line between the banded cementum and the unbanded dentine. If in doubt locate a prominent translucent band and work toward the pulp cavity following alternate transparent and opaque bands until they cease.

4. In most cases an opaque band will be found next to the cemento-dentine juncture. Usually this band is thinner than succeeding opaque bands, sometimes showing the merest trace of a line. Occasionally it is preceded by some translucent material. This band should be ignored along with the translucent bands on each side. Counting begins with the first opaque band of average size. The first opaque band plus the following (toward the outside of the tooth) translucent band constitutes the first year. Occasionally, the first opaque band appears very thick, as thick as, or thicker than succeeding opaque bands. This may be due to an obscured first translucent band so that two opaque bands are joined. Without some care, one year may be lost. Usually scanning up and down the bands in question will reveal the lost translucent band.

5. If the count of bands ends with an opaque band, add one-half year to the age. Assume these all end with ½ whether this is seen or not.

6. Occasionally opaque bands will be found divided along their length. When this splitting occurs near the middle of the tooth and counts at both ends agree, the split should be ignored. Where splitting occurs toward the ends of the tooth, the splits should be counted. In other words, the numbers of bands toward the ends of the tooth are used in each case.