Bomb dating, age validation and quality control of age determinations of monodontids and other marine mammals

Steven E. Campana[1] and Robert EA Stewart[2*]

[1]*Bedford Institute of Oceanography, Fisheries and Oceans Canada, Dartmouth, NS Canada B2Y 4A2*
[2]*Fisheries and Oceans Canada, Freshwater Institute, Winnipeg, Manitoba Canada R3T 2N6*
[*]*Present Address: 1218 Marchand Road, Howden, Manitoba, R5A 1J6*

ABSTRACT

Methods for confirming the accuracy of age determination methods are reasonably well established in fishes, but the millions of routine age determinations which take place every year require their own quality control protocols. In contrast, methods for ensuring accuracy in age determination of monodontids and other marine mammals are still being developed. Here we review the basis and application of bomb radiocarbon to marine mammal age validation, highlighting its value for providing unambiguous estimates of age for belugas and other long-lived animals which form growth bands. Bomb radiocarbon is particularly useful for marine mammals, given that the age of an individual animal can be determined to within ±1-3 years, as long as it was alive during the 1960s. However, ongoing age determinations require careful monitoring to ensure that age interpretations remain consistent across ages and through time. Quality control protocols using reference collections of ageing material, in conjunction with age bias plots and measures of precision, are capable of detecting virtually all of the systematic ageing errors that often occur once age determinations of an animal become routine.

INTRODUCTION

Age information forms the basis for calculations of growth rate, mortality rate, and productivity, ranking it among the most influential of biological variables. Calculations as simple as growth rate, or as complex as virtual population analysis, all require age data, since any rate calculation requires an age or elapsed time term. In the case of fishes, ages are usually estimated using counts of annual growth increments found in otoliths, scales, fin rays, spines, vertebrae and bones (Campana 2001). Marine mammals lack otoliths, scales, and fin rays, requiring the preparation of a tooth, bone or ear plug for direct age determination.

Age determination in marine mammals is often more challenging than it is in teleosts, in part because of the absence of otoliths in the former, and in part because the number of teleosts that are aged annually is thousands of times larger than is the case for marine mammals (Campana and Thorrold

2001). As a result, the implementation of modern age determination methods for marine mammals in general, and monodontids in particular, has tended to lag well behind that of teleosts. Ages of monodontids such as the beluga (*Delphinapterus leucas*) have typically been estimated by counting growth layers or bands (often referred to as growth layer groups, or GLGs) in longitudinal sections of the teeth. The formation of annual growth bands is now known to be ubiquitous among vertebrates (Campana 2001). However, the deposition rate of growth bands in beluga teeth was previously interpreted as semi-annual, with two GLGs representing one year of growth (Sergeant 1959). Subsequent study of beluga growth sometimes suggested annual GLG formation in the teeth, but was more often inconclusive, whether based on teeth from wild-born beluga held in captivity (Brodie 1982, Heide-Jørgensen *et al*. 1994, Hohn and Lockyer 2001) or the use of tetracycline marks as a dated chemical marker (Johnston *et al*. 1987, Brodie *et al*. 1990, Hohn and Lockyer 2001). It wasn't until the development of bomb radiocarbon as a dating tool that conclusive demonstration of annual growth band formation in beluga teeth became possible (Stewart *et al*. 2006).

In this paper, we begin by briefly reviewing the most plausible options for confirming (or validating) the age of belugas and other marine mammals. We then describe the use of bomb radiocarbon as a powerful tool for validating the age of wild, long-lived marine mammals, including belugas. We conclude by summarizing the steps required to ensure continued accuracy and precision in the age determination of monodontids and other marine mammals, as drawn from the extensive experience and history of fish ageing laboratories.

AGE VALIDATION METHODS FOR ENSURING ACCURACY

There are many high-profile examples in the scientific literature of age readers with years of experience, and the ability to provide extremely consistent replicate age readings, that were subsequently shown to be incorrect (Campana 2001). Consistency implies excellent precision, and although precision is admirable and is often the sign of a good age reader, it is all too possible to be consistently inaccurate (Svedang *et al*. 1998). Consistency does not imply accuracy. Accuracy indicates the reader is providing the correct age (on average), even if the readings are not particularly precise. Assessing accuracy usually requires some objective and independent means of determining the age of an organism. The process of confirming that accuracy is called age validation (Beamish and McFarlane 1983).

A variety of methods exist through which age interpretations can be validated. Although the distinction has often been blurred in the literature, methods can be classified as either validating absolute age, validating the periodicity of growth increment formation, or of corroborating (but not validating) an existing set of age estimates. Campana (2001) provides a complete review of age validation methods, some of which are further discussed here as applicable to marine mammals and specifically monodontids.

Age validation methods suitable for marine mammals can be ranked in descending order of rigor as: 1) release of known age and marked animals into the wild; 2) bomb radiocarbon dating; 3) mark-recapture of chemically-tagged wild animals (typically older adults of unknown age); 4) ageing of discrete length modes; 5) marginal increment analysis; and 6) rearing in captivity. A complete discussion of each of these methods is presented elsewhere in this volume; however, it is only the first three approaches that provide rigorous confirmation of age interpretation from marine mammalian teeth. As an example of the efficacy of recapturing known-age mammals, harp seals (*Pagophilus groenlandicus*) that were tagged as pups provided robust age determinations because the age of the animal was known without error (Frie *et al*. 2012). Bomb radiocarbon assays have also provided age estimates for individual belugas, but with an error on the order of ±2 years (Stewart *et al*. 2006). Bomb radiocarbon dating in general, and with beluga in particular, is discussed in more detail below. In contrast, recapture of chemically-tagged animals can accurately validate the formation of annual growth bands formed after chemical tagging (i.e. Lieberman 1993), but cannot validate the full age of the animal; for example, recapture after two years would allow the validation of only the two outermost growth bands, and only when those bands can be readily distinguished from the margin. Annual growth bands prior to the first capture mark must be assumed. The remaining age validation methods listed above (4-6) are theoretically applicable to marine mammals, but have never been applied effectively. In particular, examination of teeth from belugas reared in captivity is of questionable value, given that captive animals lack annual migratory movement and would likely exhibit different growth patterns relative to those in the wild.

Bomb radiocarbon for age validation

Bomb radiocarbon derived from atmospheric nuclear testing provides one of the best age validation approaches available for virtually any long-lived organism that forms growth bands, whether it be in trees (Worbes and Junk 1989), fish (Campana 1997), sharks (Francis *et al*. 2007), bivalves (Weidman and Jones 1993), corals (Druffel and Linick 1978), humans

(Spalding *et al*. 2005) or marine mammals (Tauber 1979, Bada *et al*. 1987, Stewart *et al*. 2006). The onset of nuclear testing in the mid-1950s resulted in an abrupt increase in atmospheric ^{14}C, which was soon incorporated into all organisms that were growing at the time. Thus the period is analogous to a large-scale chemical tagging experiment, wherein all growth bands deposited before 1956 contain only natural, low-level ^{14}C levels and all those formed after about 1968 contain elevated levels (up to 2 times natural ^{14}C levels). This measure of change in ^{14}C levels is typically presented as delta ^{14}C (Δ^{14}C; Stuiver and Polach 1977). Growth bands formed in the transition period (typically 1956 to 1968) contain intermediate and increasing Δ^{14}C levels. As a result, the interpretation of the annual Δ^{14}C time sequence (chronology) in growth bands isolated from tooth sections is relatively simple; the growth band Δ^{14}C chronology spanning the period of about 1956-1968 should match other regional Δ^{14}C reference chronologies, as long as the age assignments based on growth band counts (= years of growth band formation) are correct. Any under-ageing would phase shift the growth band ^{14}C chronology towards more recent years, while over-ageing would phase shift it towards earlier years (producing an apparent bomb ^{14}C signal prior to the start of any nuclear testing).

In the case of beluga teeth and bomb radiocarbon, the assumption that two growth bands were formed each year led to a 30-year offset of the Δ^{14}C chronology relative to the regional Δ^{14}C reference chronology. In contrast, the interpretation of one growth band per year and an age of 60 years led to an excellent correspondence between the tooth chronology and the reference chronology, thus validating the age of 60 years (Stewart *et al.* 2006). Such a shift was easily detected. Sample contamination with material of more recent origin can only increase the Δ^{14}C value, not decrease it. Thus the growth band Δ^{14}C value sets a minimum age to the sample, and the years 1956-1968 become the most sensitive years for Δ^{14}C-based ageing. Hence, for marine mammals born during this period, bomb radiocarbon dating can be used to confirm the accuracy of more traditional ageing approaches with an accuracy of approximately ± 1-3 years; the discriminatory power of samples from organisms born before or after this period is more than an order of magnitude lower because of the loss of time specificity (there is little change in natural radiocarbon levels prior to nuclear testing, while post-bomb values tend to remain high for decades) (Campana 2001).

An additional consideration for using bomb radiocarbon dating to validate age is the use of an appropriate environmental Δ^{14}C record as a temporal reference. Because the ^{14}C signal recorded in deep-sea and freshwater environments can be different from that of surface marine waters (deep-sea = delayed and attenuated; freshwater = advanced and enhanced; Broecker

and Peng 1982, Campana and Jones 1998), reference ^{14}C chronologies appropriate to the organism's environment during the period of growth band formation must be used. In addition, growth bands from non-carbonate materials (such as teeth and vertebrae) tend to be lagged by 1-2 years relative to carbonate chronologies (such as those from otoliths), because carbonate structures are usually formed from carbon taken directly from the water (in the form of dissolved inorganic carbon), whereas teeth and vertebrae obtain their carbon from food eaten by the host animal. However, the radiocarbon chronologies are usually not species-specific, which means that reference chronologies from one area and species can often be applied to another area and species. Indeed, there is relatively little variation in the timing of the initial increase in known-age (i.e. reference) ^{14}C chronologies throughout the world (Fig. 1). The global variation in ^{14}C chronologies that does exist is largely confined to the period after the Δ^{14}C rise, and can be attributed to different water mixing rates. That is, water masses in which surface waters (exposed to the atmospheric bomb ^{14}C signal) mix rapidly with deep, ^{14}C-depleted water masses (not exposed to the bomb signal), tend to show a post-bomb peak that is considerably lower and more variable (e.g. Andrews et al. 2013) than water masses where no such dilution occurs (e.g. Andrews et al. 2011). However, the year in which the bomb signal first becomes apparent (e.g. >10% above pre-bomb levels (Campana *et al*. 2008) or the first sample to exceed 2 SD of the pre-bomb mean (Kerr et al. 2006) is relatively consistent around the world in stratified waters, and thus serves as a very stable dated marker (Fig. 1). For this reason, the year of initial increase is usually considered to be the most important signal in a radiocarbon chronology; peak ^{14}C values tend to be far less informative and typically reflect atmospheric diffusion and water-mixing rates after the bomb signal first appeared in the marine environment.

In general, the bomb radiocarbon method for age validation is not well suited to studies of short-lived (< 5 yr) species (but see Melvin and Campana 2010), or when the presumed birth dates do not span the period prior to the 1960s, or in environments where appropriate reference chronologies are not available (such as the south polar region). On the other hand, the low radioactive decay rate of ^{14}C (half-life of 5730 yr) indicates that both archived and recent collections are appropriate for a Δ^{14}C assay. For example, a sample from the innermost growth band of a 50-yr old beluga tooth collected in 2010 would be just as ideal for age validation as that of the innermost growth band of an archived 10-yr old beluga tooth collected in 1970. Bomb radiocarbon assays of marine mammal teeth have an advantage over those of teleost otoliths, in that individual growth bands are often large, and thus can be sampled through micromilling in sufficient weight from any location in the growth sequence, and need not be restricted to the innermost growth band. In addition, radiocarbon chronologies can be
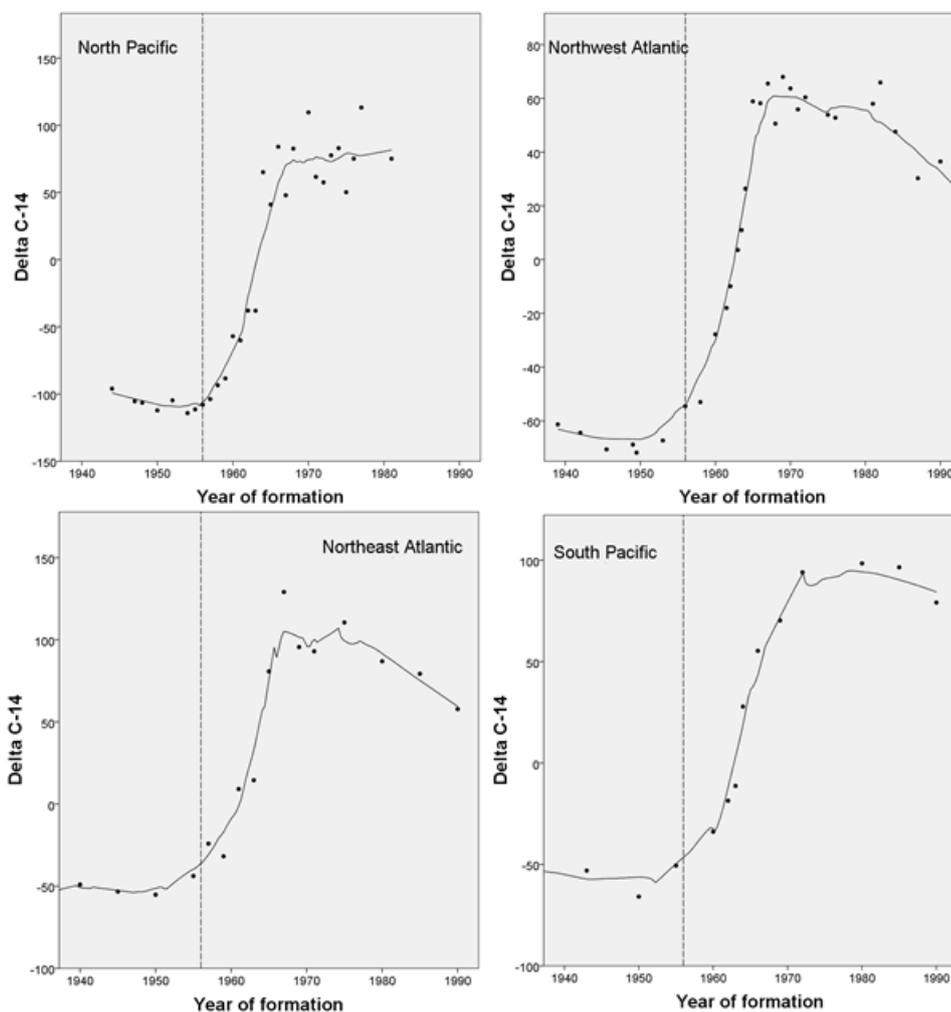
Fig. 1. A common characteristic of bomb radiocarbon chronologies from all over the world is the nearly synchronous increase from background levels around 1956 (vertical line). Thus the year of initial increase in radiocarbon serves as a dated marker in all growth band sequences. In contrast, the asymptotic (post-bomb) radiocarbon level can vary widely among water bodies due to differences in water mixing (= dilution) rate, and thus is not particularly useful as a marker. Fitted line represents LOESS curve.

developed from assays of multiple growth bands from a single individual (e.g. Stewart *et al*. 2006).

Age validation studies based on bomb radiocarbon dating are often more rapid and equally cost-effective compared to alternative methods. Despite the fact that individual assays are expensive ($500-$1000 per sample, not including the cost of growth band extraction), relatively few samples are required for age validation and processing time is measured in weeks rather than years. In contrast, validation studies using, for example, chemical tag-recapture often require substantial tagging logistics and expenses, followed by one to several years of recapture effort and rewards to fishermen.

To elaborate on how bomb radiocarbon was used to validate the age of an individual beluga whale tooth, a tooth that was sectioned and aged using conventional methods is provided to contrast age estimation methods (Fig. 2). The initial assumption was that two GLG's were formed each year (Sergeant 1959). The alternate assumption, and one that is more consistent with modern science, is that one GLG forms per year (Stewart *et al.* 2006). By counting the GLGs back from the growing edge of the tooth (corresponding to the year of death), a presumed year of formation can be assigned to each GLG (or pair of GLGs), corresponding to a year of formation. The radiocarbon assay from any single GLG can then be plotted using the year of formation (as determined from the GLG count) on the X axis. Normally, this would be done with several assays from the same or multiple teeth, so as to span a multi-year period. If the resulting time series matches that of the known-age (reference) chronology from the same region, the age and year assignments based on the GLGs must have been correct. More specifically, if the initial year of increase in the tooth chronology is similar to that of the reference chronology, the GLG interpretation must have been correct. In this example (Fig. 2B), the tooth chronology starts to increase around 1958, very soon after it increased in the fish- and coral-based reference chronology, indicating that the single GLG interpretation was correct. In contrast, GLG interpretations assuming two bands per year would produce an age half that of the original GLG interpretation, and a corresponding shift in the presumed year of formation for each pair of GLGs. When the radiocarbon assays from each GLG pair are then plotted with their new presumed year of formation, there is a huge shift in the tooth chronology, such that it appears to begin to increase in the mid-1970s (Fig. 2C). Clearly, the correspondence between the tooth chronology based on 2 GLGs per year and the reference chronology was unacceptably poor, thus rejecting the hypothesis that two GLGs form each year. Although there are statistical tests available to confirm the superior fit of one ^{14}C chronology over another (Francis *et al*. 2010), visual comparisons of fits such as those in Figure 1 are usually fairly obvious.
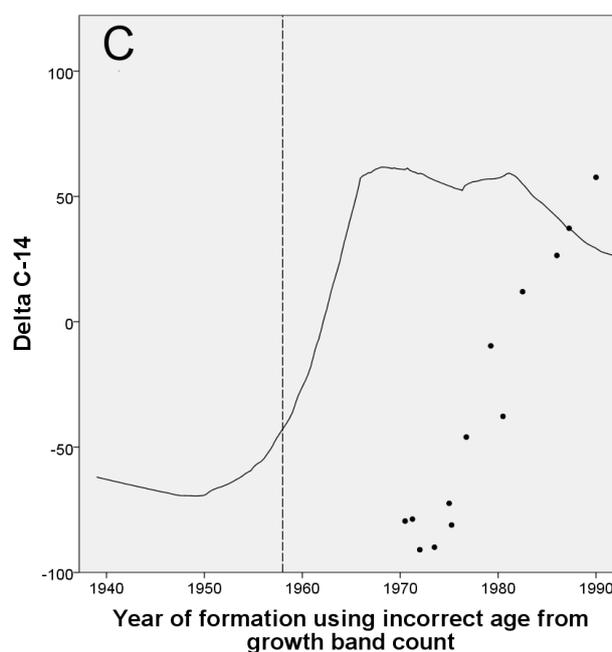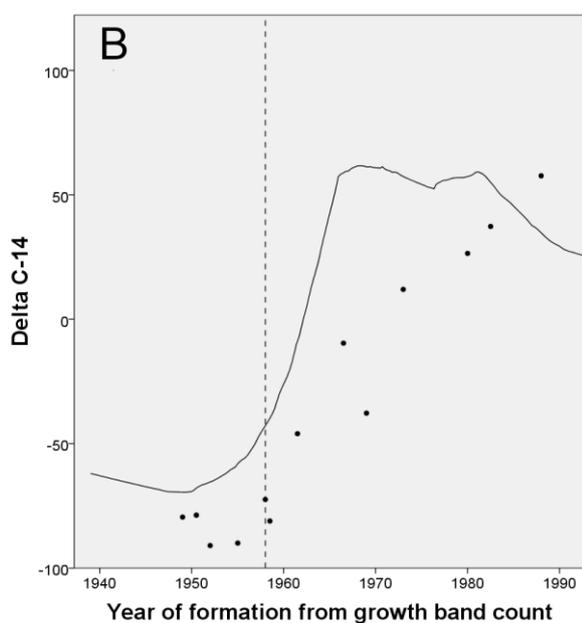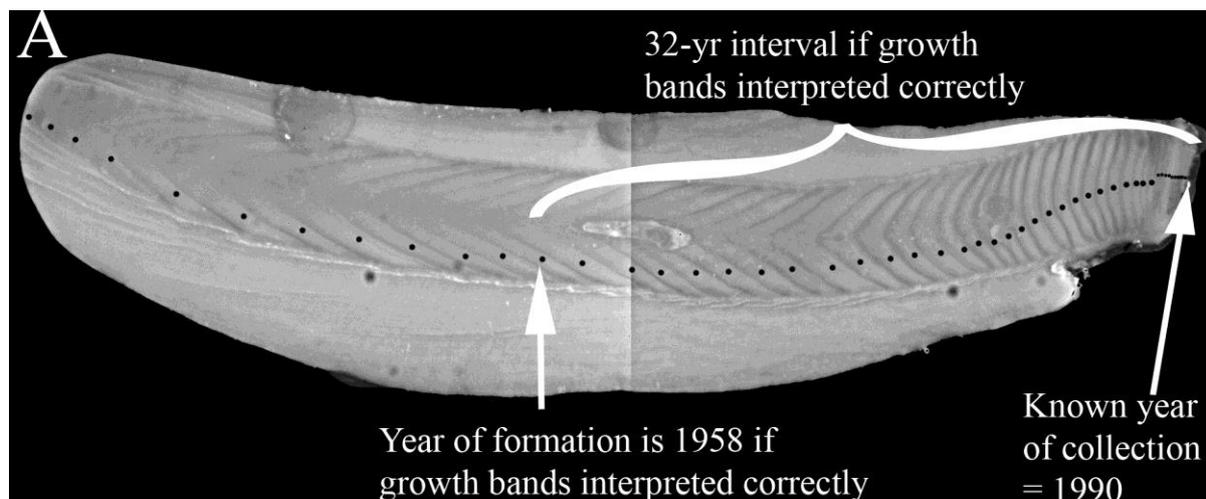
Fig. 2. (A) Longitudinal section of a beluga tooth showing presumed annual growth bands. Dates of formation can be determined by counting the growth bands in from the growing edge (right or pulp edge in this image), which corresponds to the year of collection. (B) Bomb radiocarbon assays of individual growth bands in the tooth shown begin to increase around 1958 (vertical dashed line) in synchrony with the marine reference chronology (solid line) if the growth bands have been aged and counted correctly (i.e. one GLG per year), but (C) are greatly offset from the marine reference chronology if the growth bands have been interpreted as forming twice per year (2 GLG per year). Years shown in this example are hypothetical.

REFERENCE COLLECTIONS AND QUALITY CONTROL OF AGE READINGS

There are four steps leading to the development and continued success of an ageing program: 1) development of an ageing method; 2) age validation; 3) preparation of a reference collection; and 4) quality control (QC) monitoring. In the case of beluga ageing, the use of tooth longitudinal sections is the ageing method, while bomb radiocarbon dating was the method used to validate the accuracy of the method. The latter two items—the reference collection and the quality control monitoring—are the steps required to ensure that subsequent ageing of belugas remains accurate, and that tooth interpretations do not change over time, perhaps as a result of a change in personnel who age them. More specifically, quality control monitoring can track ageing consistency through time, under the previously tested and confirmed assumption that the method is accurate. As noted by Campana (2001), the monitoring process ensures 1) that the age interpretations of individual age readers do not 'drift' through time, introducing bias relative to earlier determinations; and 2) that the age interpretations by different readers are comparable. Such a protocol monitors both relative accuracy and precision at regular intervals, and is completely analogous to quality control protocols in a manufacturing process. Integral to the quality control process is the reference collection and two statistical monitoring tools discussed later: age bias plots and the CV.

Reference collections of ageing structures are important elements of an ongoing ageing program. Ideally, a reference collection is a group of prepared and aged structures of known or consensus-derived ages and representative of all factors that might reasonably be expected to influence the appearance or relative size of the growth bands. A list of such factors might include all combinations of age, sex, season, and source of collection, spanning the entire organism length range, a representative sample of the geographic range, and several collection years. The primary role of the reference collection is to monitor ageing consistency over both the short and long term, as well as among age readers. The collection is particularly important for tests that may reveal long-term drift in age interpretation, something that cannot be detected through simple re-ageing of samples from the previous year, or through use of a secondary age reader. A second role of the reference collection is for training purposes; a representative subsample of the collection can be imaged and annotated, thus simplifying the training of new age readers and ensuring consistency in the type of structures which are interpreted as growth bands. It is important to note that ideal reference collections are rare. It is far more important to have something—anything—than to have nothing. A collection of 200 teeth or

other ageing structures is good, and can be added to through time, but again, a few dozen tooth section images is better than nothing.

Once assembled, the reference collection can be sent out for ageing as part of an exchange program, either physically or in the form of digital images. The preparation of digital images ensures long-term availability, facilitates exchanges with other laboratories, and simplifies the training of new age readers (see NAMMCO 2012 for caveats and recommendations). The use of annotated digital image 'layers' (*sensu* Photoshop), which can be toggled off and on, allows the image to be interpreted with or without the annotation and facilitates blind age comparisons in training exercises.

For quality control monitoring, a subsample of the reference collection is intermixed with a subsample of recently-aged samples (a production ageing subsample) and is then aged without the age reader knowing which samples come from the reference collection. An age bias graph comparing test versus reference ages for the reference structures would confirm long-term ageing consistency, while a separate age bias graph comparing test versus original ages for the production subsample would insure consistency between the most recent production run and the QC test. If both tests indicate lack of bias, the same ageing criteria must have been used for both reference and production samples. The CV of original and new ages provides the measure of ageing precision. The combination of the age bias graphs and CV is sufficient to detect almost all sources of ageing error.

The age bias plot is the primary tool for assessing bias, which is defined as a systematic difference between two age readers or ageing methods. It is the ideal tool for detecting under- or over-ageing of one age reader relative to another, even if the ageing error is restricted to the youngest or oldest animals (Fig. 3). Ideally, age bias plots are prepared when the age reader compares current readings against a known-age reference collection. When this is done, the age bias plot becomes a check on the accuracy of the age reader. When known ages are not available, the most reliable set of ages is used as the reference age on the X axis. However, only relative accuracy is being assessed at this point, since neither set of age readings is known to be correct. When two age readers are being compared, or when the comparison is between two ageing structures or methods, the age bias plot can only reveal a systematic difference. For example, Reader 2 in Figure 3B is under-ageing specimens relative to the reference age, but it is possible that Reader 2 is correct and that the reference age is too high.

It is important to note that the interpretation of the age bias plot is in terms of overall patterns, not at an individual age. The intent is to detect consistent deviations from the 1:1 line, not to examine the deviation of any one age

from the line. For example, in Figure 3A there are some points that lie above the 1:1 line and some points that lie below it; some mean values are statistically different from the 1:1 line (as indicated by the extent of the error bar), and others that are not. But there is no overall pattern, and therefore no bias. Conversely, there may be a trend. In the example of Figure 3B, several continuous ages (ages 4-6) are about 1 yr above the line, and ages 8 and older are all increasingly below the line. This is a more serious type of bias, since it indicates that one of the two age readers has changed their age interpretation criteria relative to the other age reader. The easiest type of age bias to deal with occurs when a reader (Reader 3 in Fig. 3C) consistently counts 1-2 extra growth bands relative to the reference age. This type of bias usually occurs when one reader is counting the edge (or a first annulus) and the other reader is not. A brief comparison of annotated images is usually sufficient to remove of this type of bias.

Ageing precision refers to the reproducibility or consistency of repeated age determinations on a given structure, whether or not those age readings are accurate. It is not unusual for inaccurate age readings to be highly reproducible (in other words, precisely wrong). Therefore, precision cannot be used as a proxy for accuracy. Nevertheless, a measure of precision is a valuable means of assessing the relative ease of estimating the age of a particular structure, of assessing the reproducibility of an individual's age determinations, or of comparing the skill level of one age reader relative to that of others.

There are two widely used and statistically robust measures of ageing precision: 1) average percent error (APE), and 2) coefficient of variation (CV). Although percent agreement is the traditional index of ageing precision, it varies widely both among species and among ages within a species. For example, 90% agreement to within one year between two age readers would represent poor precision if there were only 3 year classes in the population. In contrast, 90% agreement to within one year would represent excellent precision for beluga, given its 60-yr longevity. Therefore, there is little reason to recommend the use of percent agreement when more robust and easily-calculated measures of precision are readily available.

The coefficient of variation (CV), expressed as the ratio of the standard deviation over the mean, is the most widely used measure of precision, and can be written as:
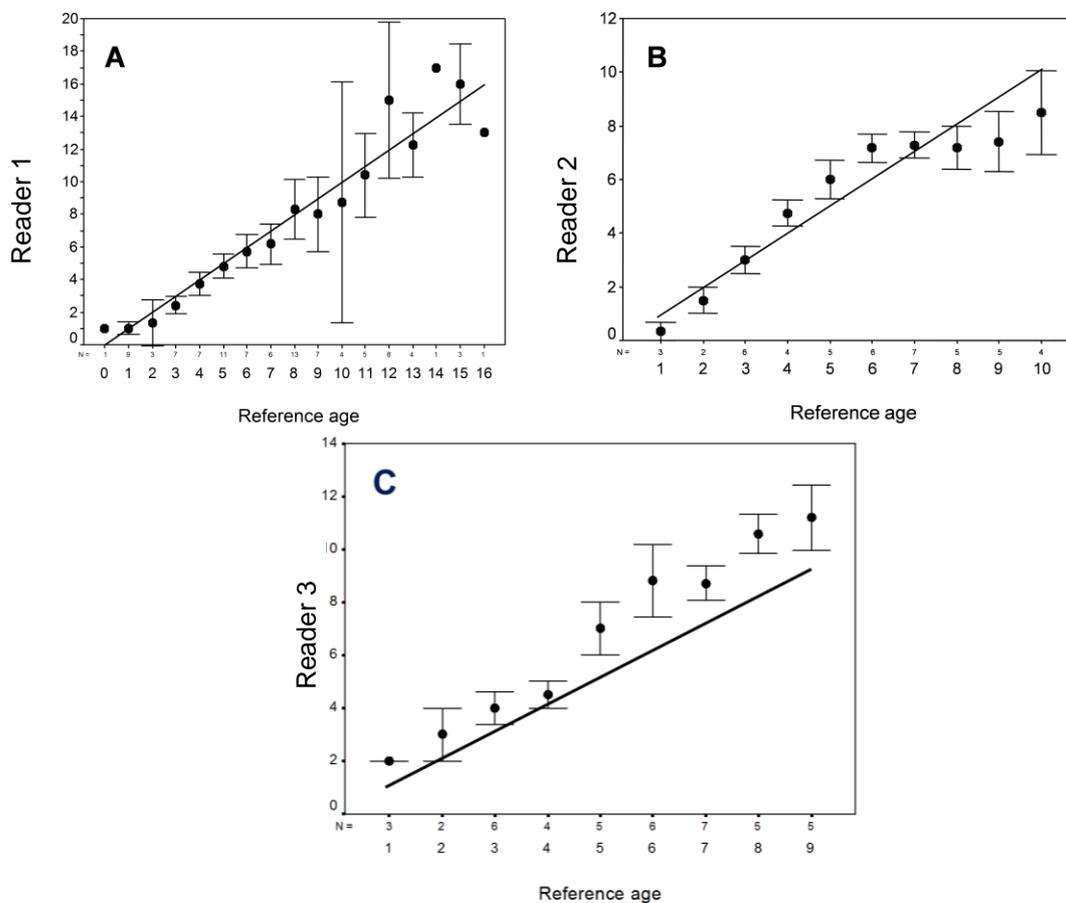
Fig. 3. Examples of age bias plots where no bias is present (A) and where bias is present (B and C). In (B), Reader 2 has over-aged ages 4-6 but under-aged ages 8-10. In (C), Reader 3 has consistently counted 1-2 extra growth bands compared to the Reference age. Each error bar represents the 95% confidence interval about the mean age assigned to all samples of a given age by a second age reader (Reference age, known or assumed to be correct). The 1:1 equivalence (solid line) is also indicated. Numbers plotted below symbols are the sample size at each age.

$$CV_j = \frac{\sqrt{\sum_{i=1}^{R} \frac{(X_{ij} - X_j)^2}{R-1}}}{X_j}$$

where CV_j is the age precision estimate for the *j*th animal. The CV is calculated across all age readings for each animal, and is usually averaged across animals to produce a mean CV.

The average percent error (APE), is defined as:

$$APE_j = \frac{1}{R} \sum_{i=1}^{R} \frac{|X_{ij} - X_j|}{X_j}$$

where X_{ij} is the ith age determination of the jth animal, X_j is the mean age estimate of the jth animal, and R is the number of times each animal is aged. When averaged across many animals, it becomes an index of average percent error.

CV and APE are mathematically related, with CV being about 40% higher than APE for any given set of ageing data (Campana 2001). All measures of precision will be artificially inflated by any bias which exists among readers, implying that bias should be dealt with before calculating precision. There is no single value of precision that can be used as a target level for ageing studies, but a CV of 5% is often used for otolith studies. CV values of more than 10% are common in studies reporting shark ages based on vertebrae (Goldman et al. 2012).

CONCLUSION

Methods for confirming the accuracy of age determination methods for monodontids and other marine mammals are more limited than those for fishes, in part because of the markedly lower numbers of animals that are available for study. However, the use of bomb radiocarbon is particularly well suited for monodontids, given their extended lifespan. In addition, the relatively large size of the growth bands in sectioned teeth or narwhal tusks makes them amenable for assaying individual growth bands, allowing individual animals to be aged with great accuracy, as long as they were alive during the 1960s. In light of the ongoing development of ageing methods for marine mammals, age validation using methods such as bomb radiocarbon dating will be required before the accuracy of the ages can be broadly accepted.

Once monodontid age determinations become commonplace, careful monitoring will be required to ensure that age interpretations remain consistent across readers and through time. Quality control protocols using reference collections of ageing material, in conjunction with age bias plots and measures of precision, are capable

of detecting virtually all of the systematic ageing errors that often occur once age determinations of an animal become routine. Use of age quality control protocols helps ensure that any observed changes in monodontid population age structure or size at age are due to real changes in the population, as opposed to artefacts resulting from changes in age interpretation.

ACKNOWLEDGEMENTS

LITERATURE CITED

Andrews AH, Kalish JM, Newman SJ, and Johnston JM (2011) Bomb radiocarbon dating of three important reef–fish species using Indo–Pacific Δ^{14}C chronologies. *Mar Freshwater Res.* 62(11):1259–1269. http://dx.doi.org/10.1071/MF11080

Andrews AH, Leaf RT, Rogers-Bennett L, Neuman M, Hawk H and Cailliet GM (2013) Bomb radiocarbon dating of the endangered white abalone (*Haliotis sorenseni*): investigations of age, growth and lifespan. Mar Freshwater Res. 64(11): 1029-1039. http://dx.doi.org/10.1071/MF13007

Bada JL, Vrolijk CD, Brown S, Druffel ERM and Hedges REM (1987) Bomb radiocarbon in metabolically inert tissues from terrestrial and marine mammals. *Geophys. Res. Lett.* 14:1065-1067. http://dx.doi.org/10.1029/GL014i010p01065

Beamish RJ and McFarlane GA (1983) The forgotten requirement for age validation in fisheries biology. *T. Am. Fish. Soc.* 112(6):735-743. http://dx.doi.org/10.1577/1548-8659(1983)112<735:TRFAV>2.0.CO;2

Brodie PF (1982) The beluga (*Delphinapterus leucas*); growth at age based on a captive specimen and a discussion of factors affecting natural mortality estimates. *Rep. Int. Whaling Comm. No*. 32: 445-447.

Brodie PF, Geraci JR and St Aubin DJ (1990) Dynamics of tooth growth in beluga whales, *Delphinapterus leucas*, and effectiveness of tetracycline as a marker for age determination. *In* Smith TG, St Aubin DJ and Geraci JR (eds.); *Advances in research on the beluga whale* Delphinapterus leucas. *Can. B. Fish. Aquat. Sci.* 224: 141-148

Campana SE (2001) Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. *J. Fish Biol.* 59(2):197-242. http://onlinelibrary.wiley.com/doi/10.1111/j.1095-8649.2001.tb00127.x/abstract

Campana SE (1997) Use of radiocarbon from nuclear fallout as a dated marker in the otoliths of haddock, *Melanogrammus aeglefinus*. *Mar. Ecol. Prog. Ser*. 150:49-56.

Campana SE and Thorrold SR (2001) Otoliths, increments and elements: keys to a comprehensive understanding of fish populations? *Can. J. Fish. Aquat. Sci*.

58(1):30-38. http://www.nrcresearchpress.com/doi/abs/10.1139/f00-177

Campana SE, Casselman JM and Jones CM (2008) Bomb radiocarbon chronologies in the Arctic, with implications for the age validation of lake trout (*Salvelinus namaycush*) and other Arctic species. *Can. J. Fish. Aquat. Sci*. 65(4):733-743. http://www.nrcresearchpress.com/doi/abs/10.1139/f08-012

Druffel EM and Linick TW (1978) Radiocarbon in annual coral rings of Florida. *Geophys. Res. Lett.* 5(11):913-916. http://dx.doi.org/10.1029/GL005i011p00913

Francis MP, Campana SE and Jones CM (2007) Age under-estimation in New Zealand porbeagle sharks (*Lamna nasus*): is there an upper limit to ages that can be determined from shark vertebrae? *Mar. Freshw. Res.* 58(1):10-23. http://dx.doi.org/10.1071/MF06069

Francis RICC, Campana SE and Neil HL (2010) Validation of fish ageing methods should involve bias estimation rather than hypothesis testing: a proposed approach for bomb radiocarbon validations. *Can. J. Fish. Aquat. Sci.* 67(9):1398-1408. http://dx.doi.org/10.1139/F10-068

Frie AK, Hammill MO, Hauksson E, Lind Y, Lockyer C, Stenman O and Svetochev V (2012) Error patterns in age estimation and tooth readability assignment of harp seals (*Pagophilus groenlandicus*): results from a transatlantic, image-based, blind-reading study using known-age animals. *ICES J. Mar. Sci*. http://dx.doi.org/10.1093/icesjms/fss169

Goldman KJ, Cailliet GM, Andrews AH, and Natanson LJ (2012) Assessing the Age and Growth of Chondrichthyan Fishes. In Carrier JC, Musick JA, and Heithaus MR (eds.) *Biology of Sharks and Their Relatives.* CRC Press LLC, New York. pp.423-451.

Grottoli AG, and Eakin CM (2007) A review of modern coral δ^{18}O and Δ^{14}C proxy records. Earth-Science Reviews 81(1-2):67-91. http://dx.doi.org/10.1016/j.earscirev.2006.10.001

Heide-Jørgensen MP, Jensen J, Larsen AH, Teilmann J and Neurohr B (1994) Age estimation of white whales (*Delphinapterus leucas*) from Greenland. *Medd. om Grøn.-Bioscience.* 39:187-193.

Hohn AA and Lockyer C (2001) Workshop to determine deposition rates of growth layers in teeth of white whales (*Delphinapterus leucas*). Draft report to NAMMCO, May 2001, SC/9/BN/4, 1-12.

Johnston DH, Joachim DG, Bachmann P, Kardong KV, Stewart REA, Dix LM, Stickland MA and Watt ID (1987) Aging furbearers using tooth structure and biomarkers. *In* Novak M, Baker JA, Obbard ME and Mallock B. (eds):*Wild Furbearer Management and Conservation in North America.* Edited by Ontario Trappers Association. Toronto, Canada. pp. 228-243

Lieberman DE (1993) Life history variables preserved in dental cementum microstructure. *Science.* 261(5125):1162-1164. http://dx.doi.org/10.1126/science.8356448

Melvin GD and Campana SE (2010) High resolution bomb dating for testing the

accuracy of age interpretations for a short-lived pelagic fish, the Atlantic herring. *Envir. Biol Fish.* 89(3-4):297-311. http://dx.doi.org/10.1007/s10641-010-9663-4

NAMMCO 2012. Report of the workshop on age estimation in beluga Beaufort, North Carolina, USA, 5-9 December 2011. In NAMMCO Annual Report, 2012, pg.476-540. Available at http://www.nammco.no/webcronize/images/Nammco/992.pdf

Passerotti MS, Carlson JK, Piercy AN and Campana SE (2010) Age validation of great hammerhead shark (*Sphyrna mokarran*), determined by bomb radiocarbon analysis. *Fish. Bull.* 108(3):346-351.

Sergeant DE (1959) Age determination in odontocete whales from dental layers. Norsk *Hvalfangst-Tidende (The Norwegian Whaling Gazette),* 48:273-288.

Spalding KL, Buchholz BA, Bergman LE, Druid H and Frisen J (2005) Forensics: Age written in teeth by nuclear tests. *Nature* 437(7057):333-334. http://dx.doi.org/10.1038/437333a

Stewart REA, Campana SE, Jones CM and Stewart BE (2006) Bomb radiocarbon dating calibrates beluga (*Delphinapterus leucas*) age estimates. *Can. J. Zool*. 84(12):1840-1852. http://dx.doi.org/10.1139/z06-182

Svedäng H, Wickström H, Reizenstein M, Holmgren K and Florenius P (1998) Accuracy and precision in eel age estimation, using otoliths of known and unknown age. *J. Fish Biol*. 53(2):456-464. http://dx.doi.org/10.1111/j.1095-8649.1998.tb00994.x

Tauber H (1979) ^{14}C activity of Arctic marine mammals. In: Berger R and Suess HE (eds.): *Radiocarbon dating.* Los Angeles: University of California Press, P. 447-452.

Weidman CR and Jones GA (1993) A shell-derived time history of bomb C-14 on Georges Bank and its Labrador Sea implications. *J. Geophys. Res*. 98(C8):14577-14588. http://dx.doi.org/10.1029/93JC00785

Worbes M and Junk WJ (1989) Dating tropical trees by means of C-14 from bomb tests. *Ecology* 70(2):503-507. http://dx.doi.org/10.2307/1937554