Environ Biol Fish (2010) 89:297–311
DOI 10.1007/s10641-010-9663-4

High resolution bomb dating for testing the accuracy of age interpretations for a short-lived pelagic fish, the Atlantic herring

Gary D. Melvin · Steven E. Campana

Received: 30 September 2009 / Accepted: 12 May 2010 / Published online: 1 June 2010

Abstract Although stock assessments of Atlantic herring (*Clupea harengus harengus*) in the northwest Atlantic have relied on tens of thousands of annual age determinations each year for more than 20 years, recent analyses have suggested that there may have been systemic ageing error. Tracking of dominant year-classes and otolith exchanges confirmed the presence of substantial ageing bias among some readers, although these approaches could not be used to identify an accurate set of ages (if any). We applied bomb radiocarbon in a high resolution dating approach targeted at the 1962 year-class to assess ageing accuracy by multiple age readers and laboratories. Although bomb radiocarbon age validation studies are typically restricted to long-lived species, the availability of archived otoliths through the 1960s and 1970s made herring an ideal candidate for an approach targeted at a single year-class, and allowed the extent of any ageing error to be quantified. The results clearly demonstrated that current age reading practices under-aged fish >6 yr of age by up to 45%, although younger fish were aged accurately. Age underestimation was due to incorrect annulus interpretation rather than non-annual otolith growth. By focusing on the period of most rapid radiocarbon increase (1962), the margin of uncertainty around the targeted bomb radiocarbon ages was reduced to 0.66 yr. This study represents the first time the bomb dating method has reached sub-annual accuracy, which makes it well suited to the age validation of short lived fish species. The use of the targeted approach has considerable promise for improving the accuracy of other bomb radiocarbon studies without the problematic assumptions associated with curve estimation and environmental effects.

Keywords Stock assessment · Ageing error · Herring · Age validation · Bomb radiocarbon · Year-class tracking

G. D. Melvin (✉)
St. Andrews Biological Station,
Fisheries and Oceans Canada,
531 Brandy Cove Road,
St. Andrews, NB E5B 2L9, Canada
e-mail: melving@mar.dfo-mpo.gc.ca

S. E. Campana
Bedford Institute of Oceanography,
Fisheries and Oceans Canada,
P.O. Box 1006, Dartmouth, NS B2Y 4A2, Canada

Introduction

Age based analytical assessment models rely on the accurate estimation of fish age to track year-classes or cohorts through the fishery (Bradford 1991; Richards et al. 1992; Campana 2001; McBride et al. 2005). Traditionally, most models have assumed ageing error to be small and random without bias or drift. The effects of random errors on model results and projection estimates are dependent upon the magnitude of the variation (Tyler et al. 1989; Kimura 1990; Bradford 1991; Dunn et al. 2002; Reeves 2003; Treble et al. 2008), but tend to produce trends in fishing mortality (F) and spawning stock biomass (SSB) similar to the true values, except on a different scale (Tyler et al. 1989; Kimura 1990; Restrepo and Power 1991; Reeves 2003; Melvin and Power 2007). However, in instances where the ageing error is biased, the impact on F and SSB, and in particular on stock projections and recruitment estimates, can be substantial, and can lead to highly inaccurate perceptions of stock status (Eklund et al. 2000; Reeves 2003; Hendrickson and Hart 2006; Melvin and Power 2006, 2007; Yule et al. 2008).

Otoliths have been used to age herring (*Clupea harengus harengus*) from stocks along the western Atlantic since the late 1950's (Parson and Winters 1972; Dery and Chenoweth 1979). Concerns over low stock abundance and high fishing mortalities, combined with inconsistencies in the Bay of Fundy–Southwest Nova Scotia region (NAFO Statistical Division 4X) catch-at-age and the age-based index of abundance, led to an independent review of herring ageing methods for the stock. This review revealed that although more than 200,000 otoliths had been aged for this stock over a period of 30 years, and despite the presence of quality control procedures, the accuracy of the age interpretations had never been confirmed. Further examination revealed that otoliths aged 20 years previously would have been interpreted very differently in recent years, leading to age discrepancies of 50–100% for the older fish (Melvin and Power 2007). In addition, recent otolith exchanges among laboratories experienced in herring ageing revealed significant, and sometimes substantial, differences (Libby et al. 2006; Sutherland et al. 2006). Although otolith exchanges amongst readers and institutes represent a valid approach to quantify ageing precision and identify potential bias (Haas and Recksiek 1995; Horn 2002), otolith exchanges provide information only on the precision of the method and the comparability among labs, not the accuracy of the technique. Nor do the exchanges indicate which, if any, of the participating labs are providing the true age. In the absence of an independent age validation study, a set of accurately-aged otoliths was not available for age calibration or correction of the catch at age matrix.

There are a variety of methods for validating the absolute age and the periodicity of growth increments of the structure used for ageing, with the recapture of chemically-tagged or known-age individuals being among the most rigorous (Campana 2001). However, herring are a fragile, locally abundant and migratory species, and thus are poor candidates for a tag-recapture age validation study. Marginal increment analysis, length frequency analysis and year-class tracking can all be effective measures for age corroboration of young, fast-growing cohorts, but are poorly suited for older fish where length and age frequency modes are blurred and incremental otolith growth is limited. For long-lived species, bomb radiocarbon has often been used to provide a dated marker against which age determination accuracy can be calibrated (Campana 1984; Kalish 1993; Campana and Jones 1998). The bomb radiocarbon method is based on the rapid increase in atmospheric and environmental radiocarbon during the 1950s and 1960s as a result of the atmospheric testing of nuclear weapons. Growth increments formed before about 1957 do not contain the enhanced levels of radiocarbon from the nuclear testing, while increments formed after about 1968 contain markedly higher levels. As a result, growth increments formed during the early 1960s contain characteristic and predictable concentrations of radiocarbon that can be used as a sensitive date marker. An important requirement of the approach is that at least some of the calcified material to be aged is derived from the sensitive 1960s time interval (Kalish 1995; Campana 2001; Kneebone et al. 2008).

Previous bomb radiocarbon age validation studies have successfully used assays of radiocarbon concentration in otolith cores to either reconstruct the entire chronology through the period of bomb signal increase, or to identify the year of initial appearance in the bomb signal (Andrews et al. 2005; Kastelle et al. 2008). The resultant uncertainty of 2–3 years has limited the application of this method to long-lived species, where an error of this duration would correspond to an (acceptable) age uncertainty of 5–7%. In this study, we take advantage of archived herring otoliths collected during the 1960s, so as to focus on the period with the most rapid increase in radiocarbon, and thus the most sensitive to bomb dating. We then target for assay the otolith cores of a single herring year-class (1962) collected through time, so as to provide a high-resolution assessment of age and ageing accuracy without any assumptions about the years of initial increase or shape of the bomb signal. The study thus represents the first application of the bomb dating method to a short-lived fish, and a resulting age validation with sub-annual accuracy.

Methods

For herring and all other organisms, the most sensitive year-classes for bomb radiocarbon dating are those hatched during the period of most rapid increase in atmospheric and marine radiocarbon: 1958 to 1965. Year to year increases in $\Delta^{14}C$ can exceed 15 during this period, which is well in excess of the uncertainty associated with any individual radiocarbon assay. In principle, maximum dating sensitivity can be achieved by focusing on a single year-class from this period (e.g., 1962), since any ageing errors which result in a year-class mis-assignment should be readily detectable in an anomalous $\Delta^{14}C$ value. Therefore, our experimental design was to first enhance the known-age reference $\Delta^{14}C$ chronology through the 1960s by assaying known-age herring otoliths of Age 0 and 1 year. We then proceeded to take advantage of the extensive herring otolith collections at the St Andrews Biological Station to follow the presumed 1962 year-class as it grew through consecutive years. That is, we used otolith collections of the 1962 year-class through Age 1 in 1963, Age 2 in 1964, etc. through to Age 10 in 1972, removing the core from representative samples in each year for radiocarbon assay. The presumed year-class was based on age assignment from annulus counts of the otolith. Any ageing errors should thus be visible as deviations from the expected $\Delta^{14}C$ value of the core.

Otolith selection

Otoliths for radiocarbon assay ($n=10$ per year) were selected from the presumed 1962 year-class from the period 1963–1972. To minimize variability due to stock mixing, selection was limited to samples from the Bay of Fundy/southwest Nova Scotia (NAFO Division 4X) collected in September of each year (Fig. 1). Otolith selection was based on the original assigned age, which may not reflect the true age of the

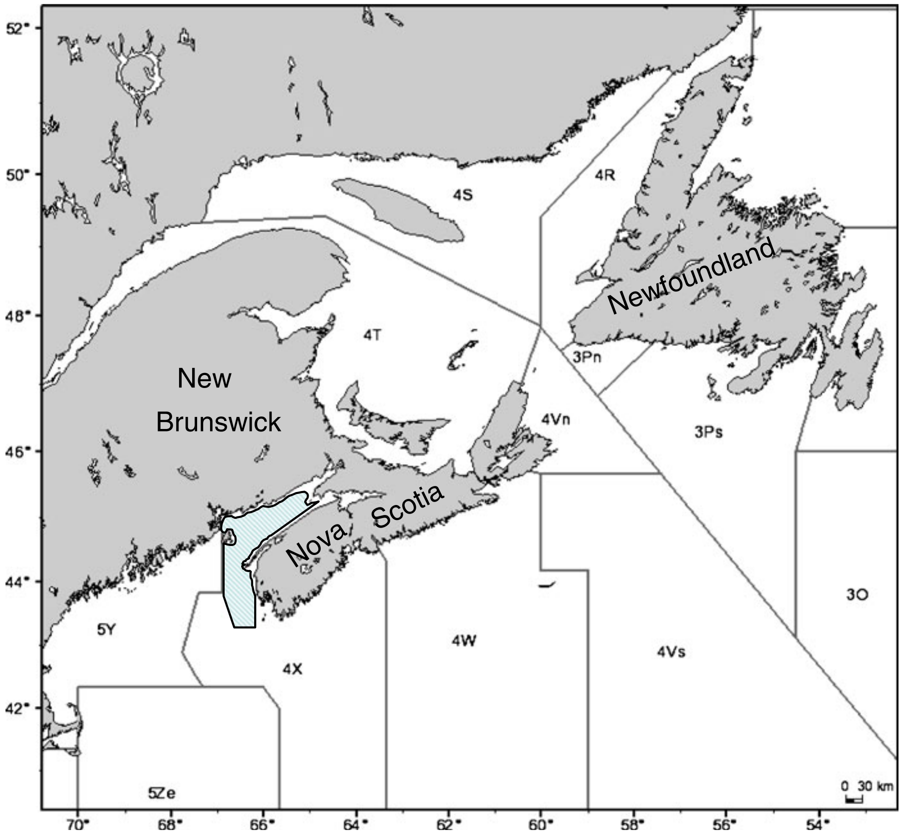

Fig. 1 Regional map and location of NAFO Statistical Divisions boundaries associated with herring spawning stocks. The shaded area represents the geographical distribution from which NAFO Division 4WX herring otoliths were collected

fish (and was tested in this study). No Age 2 otoliths were available from 1964, and only 6 Age 9 otoliths were available from 1971. An additional 10 Age 1 otoliths from 1962 were assayed for a total of 96. Before assay, the otoliths were sent to three independent laboratories for ageing by five independent and experienced age readers. Each otolith tray was randomly assigned a numerical label so that no information about fish length, date, or year of sampling was available to the readers. Month of capture was provided to assist the readers with the edge assignment. Readers were instructed to prepare and read otoliths according to standard practices at their lab.

Age interpretation

All ages were based on counts of presumed annual growth increments (annuli) visible in the intact sagittal otoliths under a binocular microscope at 16-40X magnification using reflected light. Readers immersed the otolith in water or ethanol during age interpretation. By convention, the core of the otolith (formed in the first few months of life after a fall hatch) was interpreted as the first annulus formed on Jan 1, while the translucent zone on the edge of larger otoliths was interpreted as an annulus (Dery 2005). Once all readers completed their ageing, the otoliths were sent to the Bedford Institute of Oceanography (BIO) for bomb radiocarbon assay preparation. Since the preparation for bomb radiocarbon assay is destructive by nature, all otoliths were first digitally photographed at a resolution of 2048×2048. The images were later digitally enhanced for contrast using Adobe Photoshop CS2. Identification of the second annulus for the core extraction was based on the enhanced images.

Reference $\Delta^{14}C$ chronology

The reference $\Delta^{14}C$ carbonate chronology for the Northwest Atlantic (NWA) was derived from 73 otoliths of young, known-age fish of various species whose cores were formed between 1949 and 2000 (Campana et al. 2008). Differences in $\Delta^{14}C$ among species are neither expected nor observed if the species inhabit marine waters of similar water mass characteristics. The reference chronology was supplemented by 5 samples of Age 1 herring otoliths collected in Sept-Dec of 1962 and 1963 from NAFO Division 4X. Each reference herring sample consisted of 2–6 intact sagittal pairs pooled from fish of a similar size. The mean length of the Age 1 herring was 115 mm (range of 95–135 mm), which is sufficiently distinct from adjacent length modes to give us confidence that the ages of these young fish were accurate.

To determine if the embedding or micromilling process contaminated the $\Delta^{14}C$ signature of the herring otolith, one of the above-mentioned reference samples from each of the 1962 and 1963 samples was embedded and micromilled (removing only the very edge of the otolith) and compared with matching, unprocessed samples of identical age and date of formation. The difference in $\Delta^{14}C$ value (milled-whole) was −8.1 and +2.4 for 1962 and 1963 respectively, indicating that embedding and micromilling did not appreciably alter the $\Delta^{14}C$ content of the otolith.

Age validation

Herring otoliths used in this study had been stored on plastic trays in one of two embedding resins of unknown source. All visible resin was removed from the otolith with forceps and a scalpel under a binocular microscope after first loosening with a few drops of toluene or acetone. Although the resin appeared to be completely removed, the success of the resin removal process was later assessed by assaying samples of the isolated embedding matrix and comparing it with assays of the otoliths. The mixing models used by Stewart et al. (2006) were used to test for contamination of the otolith signature by the embedding matrix.

Otolith cores for bomb radiocarbon age validation were isolated from a sagittal section of the otolith prepared by polishing. Samples were isolated from sagittal sections, rather than transverse sections, so as to maximize the amount of sample material available for assay. Herring otoliths were first embedded in a slow-drying hard epoxy (Araldite epoxy GY502 and hardener HY956 in a 5:1 weight ratio). The distal surface of the embedded otolith was then polished on Imperial lapping film (3–30 μm grit size) until all embedding medium had been removed from the area to be micromilled. Previous work with transverse sections had demonstrated that otolith growth on the distal surface was negligible after the second annulus (first year of growth). The embedded otolith was then flipped and polished to a section thickness of

0.32 mm, corresponding to the thickness of the second annulus. Otolith cores representing the first year of life (to the second annulus) were isolated from the section as a solid piece with a Merchantek computer-controlled micromilling machine using 300-μm diameter steel cutting bits and burrs. Since the mean weight of isolated core material was ~2 mg per otolith pair, otolith cores from 1–5 additional fish of the same original assigned age and similar length were extracted and pooled so as to bring the sample weight up to the minimum of 3 mg necessary for radiocarbon assay.

After sonification in Super Q water and drying, the sample was weighed to the nearest 0.1 mg in preparation for ^{14}C assay with accelerator mass spectrometry (AMS). AMS assays also provided $\delta^{13}C$ ($^{0}/_{00}$) values, which were used to correct for isotopic fractionation effects and provide information on the source of the carbon. Radiocarbon values were subsequently reported as $\Delta^{14}C$, which is the per mil ($^{0}/_{00}$) deviation of the sample from the radiocarbon concentration of 19th-century wood, corrected for sample decay prior to 1950 according to methods outlined by Stuiver and Polach (1977). The mean standard deviation of the individual radiocarbon assays was about $5^{0}/_{00}$.

The date of sample formation based on the annulus-based age was calculated as the year of fish collection minus the age span of the fish from the edge of the otolith to the midpoint of the growth increments present in the extracted core. Since not all age readers assigned the same age to each assayed otolith, dates of sample formation were calculated separately for each age reader.

A quadratic equation was used to describe the relationship between year and reference $\Delta^{14}C$ values of the otolith core between 1958 and 1965. This equation was used to predict the year of formation and subsequently the age based on the observed bomb radiocarbon assay results.

$$Y_R = a(\Delta^{14}C)^2 + b(\Delta^{14}C) + c$$
$$Y_B = Y_R + 1957$$
$$A_B = Y_S - Y_B + (A_C/2) + 1$$

where Y_R = regression year = Y_{YC} – 1957, Y_{YC} = year class, Y_B = bomb radiocarbon year of formation, Y_S = sample year, A_C = core age, and A_B = bomb-based age.

Year-class tracking

The year-class tracking study was designed to investigate the reader's ability to track a dominant, known year-class as it progressed through the fishery. Several strong year-classes have been documented for 4WX herring (Power et al. 2007); however, the 1983 year-class, which dominated the fishery for about a decade, was selected for study. A representative sample of approximately 200 otoliths per year (that included the dominant year-class) was selected from those years where the cohort was observed in the fishery (total n=1787). Trays containing 30–50 otoliths sampled from purse seine catches in September (to minimize variation in interpretation of the outer edge) from 1985 to 1993 with a presumed age range of 2–14 yr were selected at random from the archives. Based on the originally-assigned ages, the 1983 dominant year-class was clearly visible in each of the 9 years selected for the study. Importantly, the number and placement of the otoliths from the dominant year-class were randomized within each year, and were unknown to the readers.

Otolith readers in the year-class tracking study originated from the same institutes as those of the bomb radiocarbon study, but some of the individuals differed. Reader 2 did not participate in this study. All readers were instructed to prepare and read the otoliths according to their standard lab protocols, and to categorize the otoliths into three groups: readable, difficult to read, and unreadable. The percentage of otoliths successfully aged by the readers ranged from 64–99% with the more experienced readers generally reading the higher percentages. Readers were provided only with the month of sampling. No information was provided on the fish size, the dominant year class, range of years, or maximum age in the samples. Age comparisons among age readers were made with age bias plots (Campana 2001).

Results

Year-class tracking

A matched comparison among the four age readers, and between the age readers and the originally-assigned age, demonstrated considerable bias. All age readers under-aged relative to the original age assignment, with the extent of under-ageing dependent upon the reader (Fig. 2). Significant bias first appeared as early as Age 5 for one reader, and as late as Age 8 for another reader. By Age 12, the magnitude of the under-ageing bias ranged from 3.5 to 5 years, or 29–41%. This estimate of bias underestimates the actual bias, since the bomb radiocarbon assays (described in following section) indicated that even the originally-assigned ages underestimated the actual age of Age 8+ fish. The consistency amongst readers was also very poor, with none of the comparisons with the originally-assigned ages reaching 80% agreement or falling below a CV of 5%.

The yearly frequency distribution of the originally assigned ages for the randomly selected otoliths clearly tracked the 1983 year-class from age 3 in 1986 when it first showed up in the fishery, through to age 10 in 1993 (Fig. 3). Some blending may have occurred between ages in the older years, especially ages 9 and 10 in 1992 and 1993. Unfortunately, only Reader 3 was able to track the year-class clearly up to age 9, although the year-class was still visible at age 10. For all other readers it was virtually impossible to identify the year-class (or any other year-class) as strong relative to adjacent ages at any age greater than 5. In fact for two of the four readers, the 1983 year-class had completely disappeared by 1991 at age 8 (Fig. 3). This is consistent with the age bias plots and reflects the tendency to under ageing by several of the readers.

Bomb radiocarbon assays

The two media used to embed the herring otoliths in the 1960s were very similar to each other, but very different from that of the otoliths: $\delta^{13}C$ was −25.1 and −32 for the embedding medium and Diatex coating, respectively, while $\Delta^{14}C$ was −997 and −975 respectively. In contrast, the mean $\delta^{13}C$ of the herring otoliths was −3.75 (SD = 0.6) while the mean $\Delta^{14}C$ was −1.2 (SD = 19.6). Based on the mixing model and phase plot approach of Stewart et al. (2006), there was no evidence of any contamination of the otolith cores with the original embedding medium.

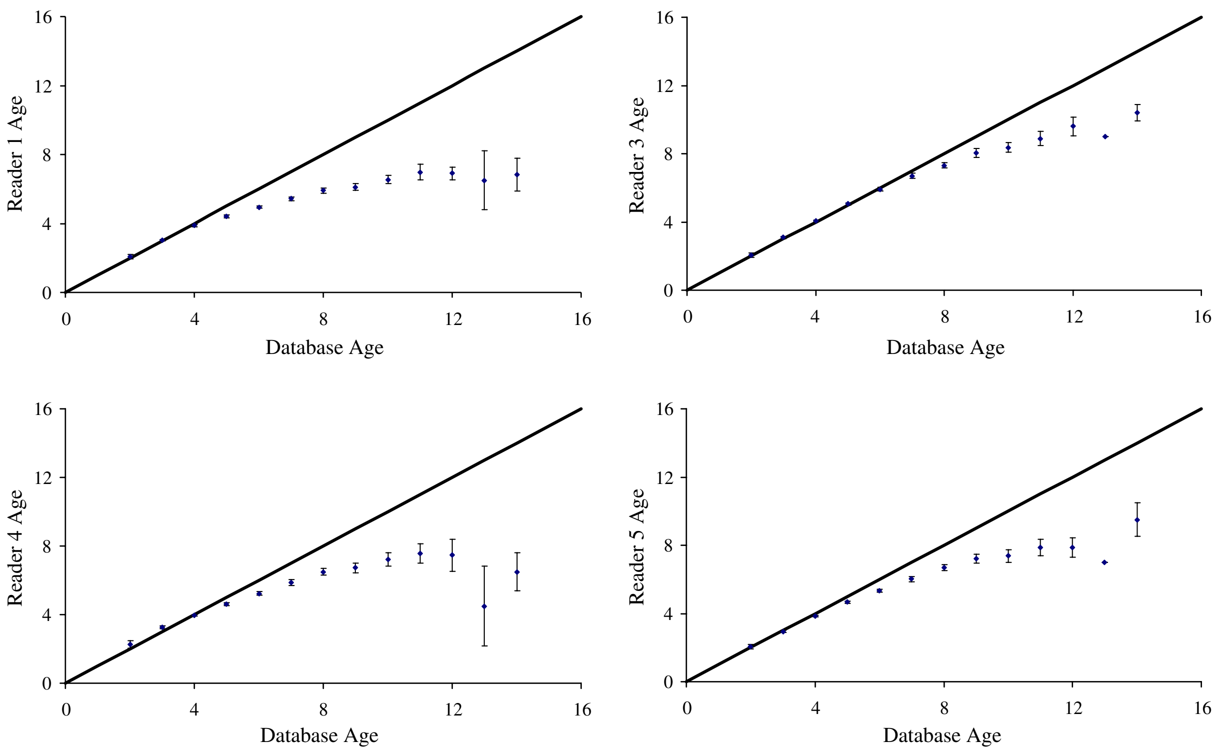

Fig. 2 Age bias plots of the originally-assigned age (database age) versus reader assigned ages for the otoliths used in the year-class tracking study. The 1:1 relationship is shown as a solid line

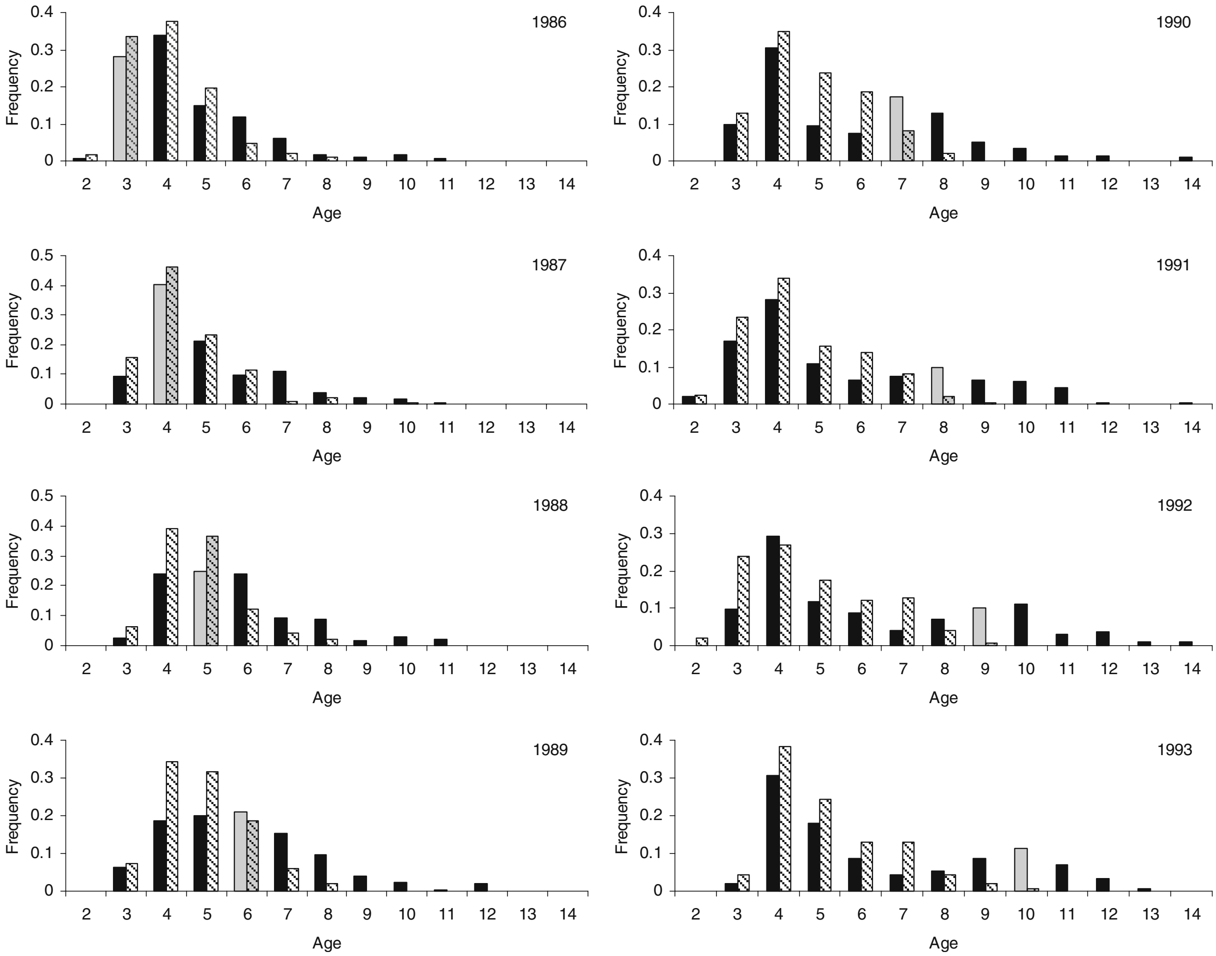

Fig 3 Age-frequency distribution standardized to the number of otoliths selected yearly at random from the database (*solid*) for the dominant year-class tracking study and as assigned by Reader 1 (*shaded*). The grey bar in each year identifies the dominant 1983 year-class

The Northwest Atlantic (NWA) reference $\Delta^{14}C$ carbonate chronology is one of several that provide a benchmark (i.e., calibration) for tracking temporal changes in bomb radiocarbon observed in fish sagittal otoliths and other calcified structures. Prior to 1958, the NWA chronology showed a relatively low and stable $\Delta^{14}C$ (Fig. 4 insert). This was followed by a period of rapid increase between 1959 and 1969 as a result of atmospheric testing, and a gradual decline after 1970. Assays of young, essentially known-age herring from 1962 and 1963 demonstrated that there was no discernable difference between the herring $\Delta^{14}C$ chronology and that of the NWA reference chronology (Fig. 4). A quadratic equation fit to the radiocarbon data between 1958 and 1965 provided a good fit to the chronology (r^2 = 0.98), and thus a robust means of estimating the mean year of core formation based on the core radiocarbon value:

$$Y_R = -0.000501\left(\Delta^{14}C\right)^2 + 0.0652\left(\Delta^{14}C\right) + 5.810$$

The mean $\Delta^{14}C$ of known-age Age 1 fish collected in 1963 (representing the 1962 year-class) was 11.6 (± 95% analytical uncertainty of 10) (Table 1). Given the mean core age of 0.9 yr, this value represents the expected radiocarbon value on the decimal year 1963.5. The presumed year of core formation of older herring otoliths was almost identical at 1963.6 (mean core age of 1.2), and thus should contain the same amount of radiocarbon as the Age 1 fish if their year of formation was accurately interpreted based on annulus counts. However, the radiocarbon content of

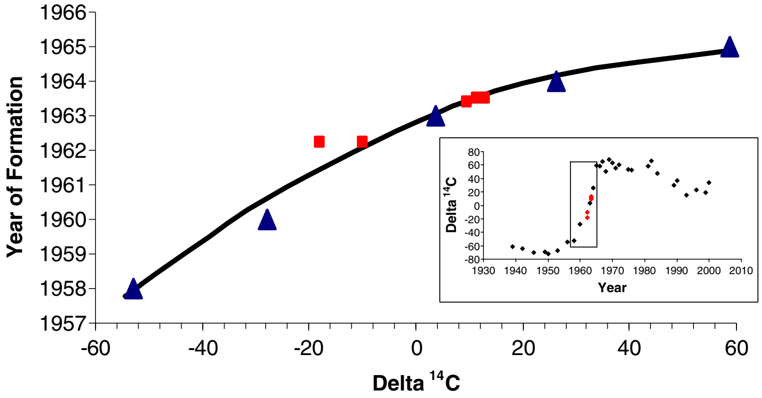

Fig. 4 Relationship of $\Delta^{14}C$ observed in young herring otoliths of known age (red) and the year of core formation relative to the Northwest Atlantic chronology (black triangles) for the years 1958 to 1965. A quadratic equation has been fit to the data (see text). The insert shows the bomb radiocarbon ($\Delta^{14}C$) reference chronology for the Northwest Atlantic (black symbols) with the rectangle high-lighting the region where the main figure was derived

the older herring otolith cores deviated from the expected value, indicating that the actual year of core formation was not that which was assumed based on the originally-assigned annulus counts (Fig. 5). For pooled herring otoliths aged originally at 5 yr and older, the observed $\Delta^{14}C$ was depleted relative to the expected value, indicating that the actual mean age was older than originally estimated.

Conversion of the observed core radiocarbon values to actual years of core formation, and then to fish age, allowed comparison of radiocarbon-based fish ages with ages based on annulus counts for each of the age readers (Table 1; Fig. 6). The annulus-based ages tended to reflect actual (radiocarbon-based) age well until at least Age 6, at which point the former tended to increasingly underestimate actual age. The magnitude and age at which underestimation first became evident varied substantially among age readers, with Reader 3 showing relatively little bias until after Age 9, and Reader 1 showing significant bias after Age 6 (Fig. 6). The extent of the age underestimation at Age 12 varied from about 1 yr for Reader 3 to about 3 yr for Reader 1. In most cases, the magnitude of the age underestimation increased with age.

Comparison of the originally-assigned (database) age with that of the radiocarbon-based age indicated that age underestimation first became evident around Age 7, and remained relatively constant at 1.5–2.0 yr for older ages (Fig. 7). In contrast, the median of the five reader ages indicated that age underestimation first became evident around Age 7, but then increased steadily with age to about 5 yr. When the analytical uncertainty (95% CI) associated with the radiocarbon assays is added to the residual plot of Fig. 7, the precision of the age underestimation estimate can be better assessed. In this case, the ageing bias is statistically significant for the older ages ($P<0.05$).

Once the correct (radiocarbon-based) age was known, it was then possible to review the digital images of the assayed otoliths and determine the correct annulus interpretation. In almost all cases, the correct interpretation of the otolith growth pattern was relatively straight forward if all of the narrow growth bands observed on the postrostral edge were interpreted as annuli, as opposed to sub-annual growth interruptions (Fig. 8). Annuli were also usually visible along the rostrum as well.

Discussion

Our results indicate that unsectioned herring otoliths can provide an accurate age estimate of the fish, but that individual reader differences in annulus interpretation can result in significant, and sometimes substantial, under-ageing. Otoliths have been the body part of choice for ageing herring in Atlantic Canada since the mid-1960's. Despite early concerns over inter- and intra-institute inconsistencies amongst age readers, no age validation study has ever been undertaken (Watson 1965; Tibbo 1970; Hunt et al.

Table 1 Summary of herring sample characteristics, pooled otoliths, mean reader age, $\delta^{13}C$ ($^{0}/_{00}$) and $\Delta\ ^{14}C$ assay results

Assay Number	Number of Pooled Otoliths	Year Sampled	Mean Length (mm)	Original Age	Mean Reader Age	$\delta\ ^{13}C$	$\Delta\ ^{14}C$	Bomb Age
15053	6	1962	105	1	1.2	−4.01	−18.0	2.0
15054	4	1962	102	1	1.2	−4.08	−9.9	1.3
15056	4	1963	119	1	1.0	−4.34	12.9	0.9
15057	5	1963	124	1	1.1	−4.71	9.7	1.1
15058	4	1963	132	1	1.0	−6.10	11.3	1.0
15069	4	1965	234	3	3.2	−3.54	20.7	2.7
15070	3	1965	236	3	3.0	−3.23	30.0	2.3
15071	3	1965	235	3	2.9	−3.61	33.3	2.2
15059	2	1967	310	5	5.7	−4.04	16.3	4.9
15060	2	1967	315	5	6.3	−3.49	−7.5	6.3
15062	3	1967	316	5	5.7	−3.44	7.9	5.3
15063	2	1967	336	5	7.2	−3.28	−13.4	6.8
15064	3	1968	310	6	5.6	−3.59	43.6	4.9
15065	3	1968	319	6	6.3	−3.77	−5.7	7.2
15066	2	1968	318	6	7.2	−3.07	−13.6	7.8
15067	2	1968	331	6	6.7	−3.33	−12.4	7.7
15077	2	1969	341	7	6.3	−3.70	−25.8	9.8
15078	3	1969	339	7	7.6	−3.93	−13.9	8.8
15079	3	1969	340	7	8.1	−3.33	−18.5	9.2
15080	2	1969	337	7	7.0	−3.63	−20.6	9.3
15081	3	1970	344	8	7.8	−3.77	25.7	7.4
15082	3	1970	349	8	8.1	−3.38	−1.6	8.9
15083	3	1970	341	8	7.1	−4.00	−8.8	9.4
15093	3	1971	356	9	7.5	−3.58	−10.0	10.5
15094	2	1971	338	9	9.2	−3.59	−11.3	10.6
15095	4	1972	360	10	7.8	−3.63	16.2	9.9
15096	3	1972	351	10	9.0	−3.27	−12.4	11.7
15097	3	1972	361	10	8.7	−3.68	−28.4	13.0

1973; Cleary et al. 1982). In the western Atlantic, herring otolith exchanges between national and international institutes ceased after 1982, and did not resume until concern was expressed about the ageing and assessment inputs for transboundary herring stocks (Overholtz et al. 2004). These concerns were amplified in 2002 when three independent research laboratories reported significantly different (biased) ages in an otolith exchange (Overholtz et al. 2004). The matched comparisons among age readers reported in this study, as well as the comparison with otoliths aged 20 yr previously, are all consistent with the presence of interpretational error leading to age bias. However, these types of comparisons do not identify the source of the bias, or indicate who (if anyone) is ageing the otoliths correctly. An age validation study, using otoliths whose age can be confirmed objectively, is required to determine ageing accuracy (Beamish and McFarlane 1983; Campana 2001).

Our study used both age corroboration and age validation techniques to determine the ageing accuracy of several age readers. Tracking of dominant year-classes through a catch at age matrix is a commonly applied, but somewhat ad hoc method, of inferring ageing accuracy (Morison et al. 1998a; Campana 2001; Kimura et al. 2006). The underlying principle is that ageing error will result in the apparent

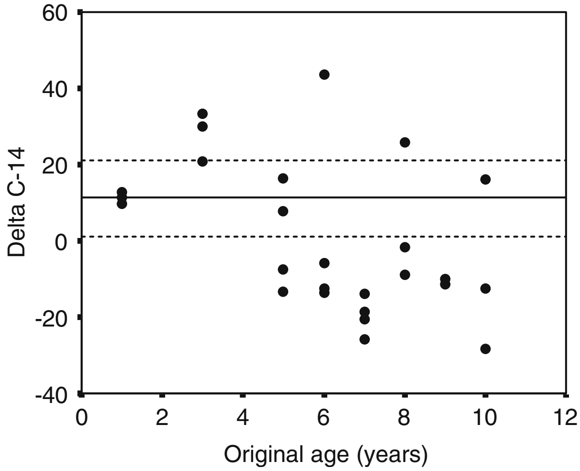

Fig. 5 Delta C-14 (Δ^{14}C) assays of herring otolith cores (mean age of 1.2 yr) assumed to have formed in 1963.6 based on the originally-assigned ages. The trend in radiocarbon values indicates that the actual year of core formation increasingly diverged from that which was assumed. Solid line shows mean radiocarbon value for Age 1 herring from the 1962 year-class. Dashed lines show the 95% confidence intervals around the mean radiocarbon value based on analytical uncertainty (SD = 5)

weakening, blending with other ages, or perhaps disappearance of the year-class at older ages, whereas correct ageing will allow the year-class to be tracked for many years. The approach does not distinguish between ageing bias (e.g., underageing) and normal, random ageing error. Our results indicated that the originally-assigned ages did a creditable job of following the strong 1983 year-class until Age 8, after which ageing error eliminated the relative dominance of the year-class. These results are completely consistent with the bomb radiocarbon assay results, which indicated that the original age assignments were more accurate than those of any of the age readers. In contrast, most age readers were unable to follow the strong year-class beyond Age 6, which is again consistent with the bomb radiocarbon results.

The advantage of the bomb radiocarbon assays was that the extent of the ageing bias could be quantified, both at older ages when ageing error was most acute, and at younger ages when the dominant year-class was still present (but perhaps diminished). This quantification indicated that both the originally-assigned and reader ages were relatively accurate until at least Age 6. After age 6 under-ageing increased to the point where the reader age under-estimated the actual age by up to 45%. In some cases almost all annuli beyond the 6th were missed, even up to age 13. The bomb radiocarbon assays also easily identified systemic ageing error (ageing bias, as opposed to random ageing error), which was not the case with the year-class tracking study. Theoretically,

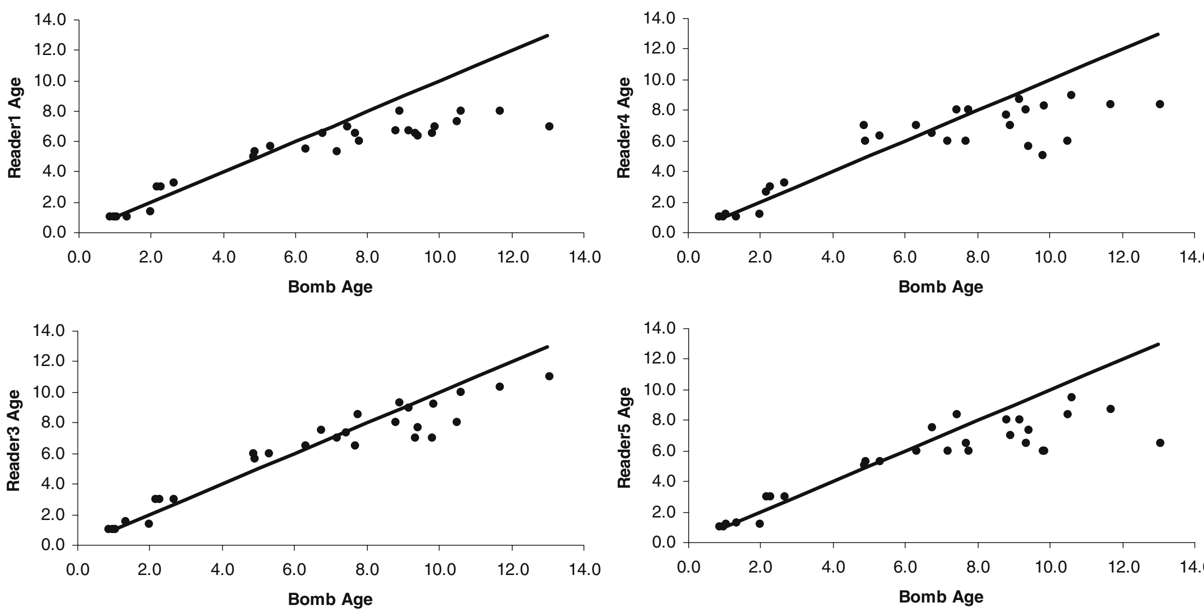

Fig. 6 Scatter plot of individual reader age for 4 of the 5 readers versus bomb radiocarbon (actual) mean age. The diagonal line represents the 1:1 age ratio. Reader 2 results were similar to those of Reader 3 and are not shown

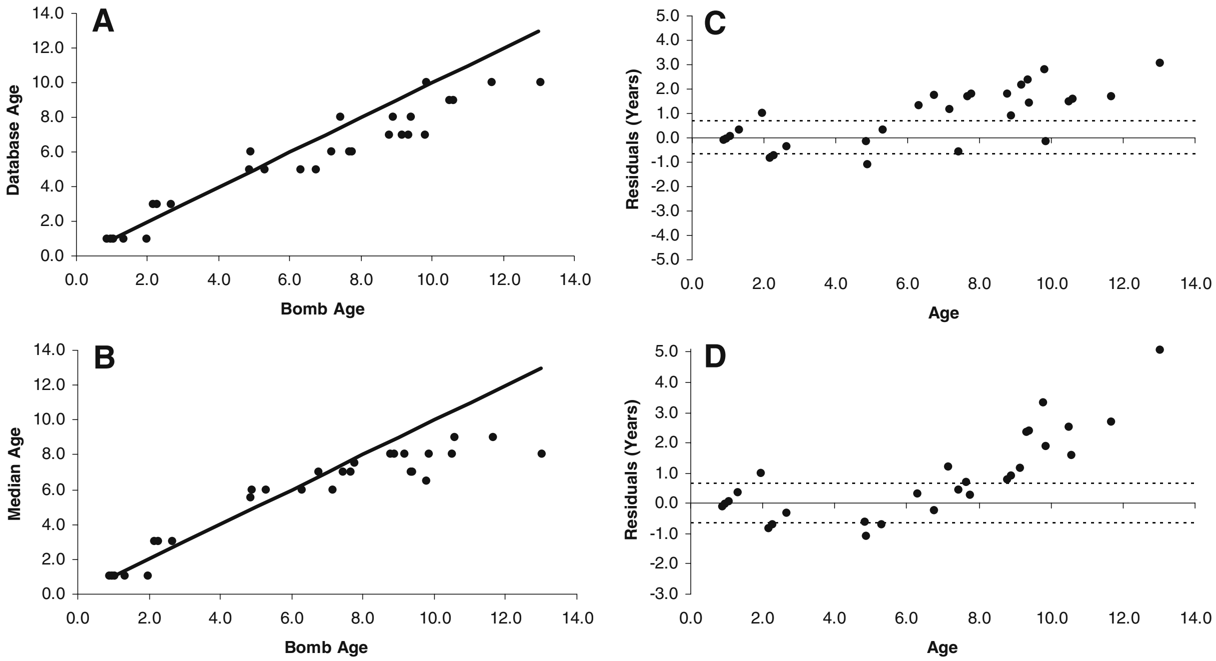

Fig. 7 Scatter plot of the originally-assigned (database) age versus bomb radiocarbon assay age **a** and median reader age versus bomb radiocarbon assay age **b** with corresponding residuals plots (C and D respectively) by age. Dotted lines in C and D represent the 95% confidence interval based on analytical uncertainty associated with the radiocarbon assay

it may be possible to quantify ageing bias in a year-class tracking study using sophisticated statistical methods, but we are unaware of any such attempts. Nevertheless, both studies indicate that under-ageing was most evident in older fish, a result which has been commonly reported in a broad range of fish species (Francis et al. 2007; Neilson and Campana 2008; Bruch et al. 2009).

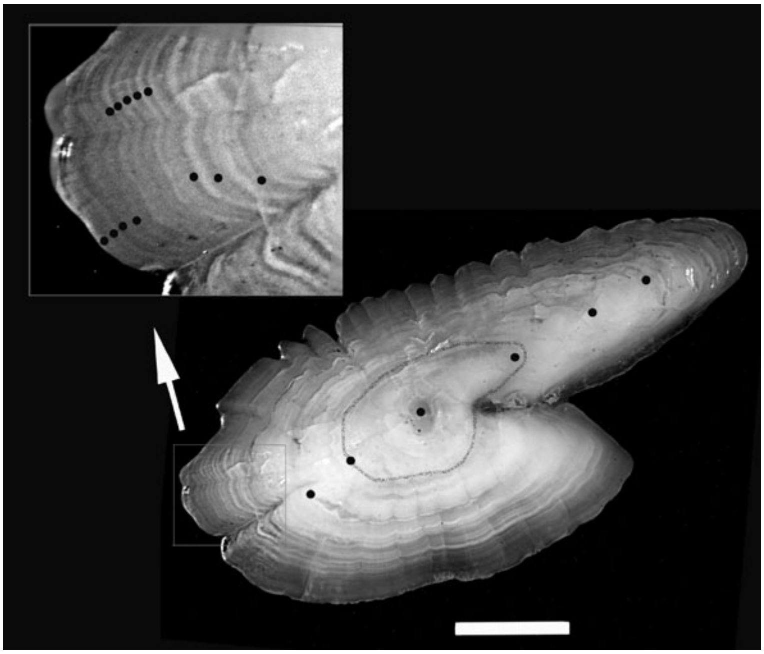

Fig. 8 Digital photo of an Age 15 herring otolith annotated to show the annuli of a relatively old fish. By convention, the core is considered to represent the first annulus. The region interior to the grey line (second annulus) was that typically micromilled for radiocarbon assay, although this particular otolith was not assayed. Scale bar = 1 mm

Although bomb radiocarbon has been used successfully to validate the age of many fish species (Andrews et al. 2007; Kastelle et al. 2008), the approach used in this study significantly improved upon both the accuracy of the age validation and the range of species' lifespans to which it could be applied. Previous studies using bomb radiocarbon have evaluated the correspondence between the reference $\Delta^{14}C$ chronology and the $\Delta^{14}C$ chronology of the test species, either through comparison of the entire chronology (Hamel et al. 2008) or the estimation of the initial year of radiocarbon increase (Campana et al. 2008). Under both approaches, precision has been limited to no less than 2–3 yr due to uncertainty in parameterizing the curve or initial year, or due to differences in the environment between the reference chronology and the species being studied. The approach taken in this study, to focus on a single year-class through time, eliminates both of these constraints: there is no longer a need to determine either the shape of the bomb curve or to estimate the initial year of increase. Rather, one need only compare the otolith core assay values of each collection year with the expected (or measured) core value of the target year-class, with the expectation that it will remain constant across sampling years if there is no ageing error. The only assumption of the method is that all of the test otolith samples come from the same cohort, thus eliminating any environmental variability. Pooling of test otoliths for a given sample also assumes that the ageing error is comparable among otoliths, although only the precision (as opposed to the accuracy) of the result would be affected if it was not. The only remaining uncertainty is the analytical uncertainty associated with the radiocarbon assay itself, plus any error associated with extraction of the otolith core. Given a typical radiocarbon assay uncertainty of ±10 (95% CI), analysis of a year-class in the period of rapidly-increasing radiocarbon (e.g., 1962) would translate into an age uncertainty of ±0.66 yr for any individual otolith. In principle, it should be possible to improve even further on this level of precision, by using larger sample sizes or pooling otoliths of a common age.

To our knowledge, this study is the first to apply the bomb radiocarbon age validation method on such a short lived species. Two factors contributed to this successful application. First, the archived otolith collection was sufficiently large that annual collections of a single year-class were available over a 10 year period, allowing a focus on a single year of core formation across time. Secondly, and perhaps most importantly, the targeted year-class represented the steepest part of the radiocarbon chronology (early 1960s), and thus the core $\Delta^{14}C$ provided the greatest discrimination between adjacent year-classes. This condition is particularly important for a short-lived species, since an uncertainty of, say, 2–3 yr, would not be helpful in trying to confirm the age of a 6-yr old fish. These two conditions will not always (or even often) be available. However, several variations on this approach are more readily accessible. For example, it is not necessary that the otolith core be formed during the 1960s, only that a growth increment containing sufficient material for micromilling be formed during that period. In addition, sequential annual collections are not required; they could be intermittent. For example, one could assay the 4th annual increment from the core in a long-lived species collected in 1978 at a presumed age of 20 yr and believed to have been hatched in 1958, and then assay the 4th increment from the core of a 30-yr old fish (the same cohort) in a collection made 10 yr later in 1988. In this example, consistency in the radiocarbon assay between the two samples would validate the interpretation and annual nature of growth increments formed between Age 20 and Age 30. Although such an application would not confirm the overall age estimate of the fish (since the periodicity of growth increments prior to Age 20 would not have been confirmed), it would validate the interpretation of the most recently formed (oldest) increments, which are often the most problematic for age determination (Beamish and McFarlane 1983). An analogous sampling design could equally be applied to a very short-lived species. Interestingly, this type of bomb radiocarbon application essentially becomes a tag-recapture study, with the radiocarbon content of the targeted growth increment becoming the tag, and the period between collections becoming the time at liberty.

The results of this study highlight an unpleasant reality: many years of ageing experience, tens of thousands of otoliths read, high repeatability within age readers, and the involvement of multiple labs in otolith exchanges, do not necessarily mean that the resulting ages are correct. Of course, otolith exchanges fill an important role in highlighting inconsistencies in age interpretations among institutes that may in turn identify an ageing problem. However, only an age validation study using objectively-determined ages can confirm ageing accuracy. In the case of Atlantic herring, the accuracy of the originally-assigned ages was actually very good for the ages making up the bulk of the catch at age matrix (up to Age 6), thus leaving the inaccuracy of the ages for the older fish undiagnosed for many years. However, institutional errors in production ageing (Morison et al. 1998b) continue to be reported (Beamish and McFarlane 1995; Campana 1997; Power et al. 2006; Bertignac and de Pontual 2007), and are probably more common than is now recognized. Fortunately, improved quality control procedures should make such errors less frequent in future (Appelberg et al. 2005). In particular, the use of reference collections based on known-age or validated material, with regular additions of new calibrated material to reduce memorization of individual otoliths, is recommended to avoid systematic ageing error (Campana 2001).

Nothing in this study suggested that intact herring otoliths do not produce reliable and interpretable annual growth increments. Indeed, the opposite was true—when the true age of the otolith was revealed after bomb radiocarbon assay, the stored digital image of the otolith invariably revealed the expected number of increments in predictable locations. Thus, it is reasonable to conclude that intact western Atlantic herring otoliths can provide accurate indicators of age at all ages, but that calibration against known-age otoliths can simplify the interpretation of the narrower outer annuli. Such is not the case with many longer-lived species, where otolith sections are required to reveal annuli in old fish (Dwyer et al. 2003).

Although the age validation of the herring otoliths is now complete, several stages remain before the entire problem has been addressed. The preparation of a reference collection containing the bomb-aged otolith images will help ensure that future age readers can calibrate their interpretations against otoliths of known age (Campana 2001; Maceina et al. 2007). However, there is no simple adjustment factor that can be applied to the previously-collected ages in order to correct for the underageing; given the variable year-class strength of herring, an age-specific correction factor would be inappropriate for year-classes of different strength. Thus the only realistic solution is to re-read a subset of the otoliths from multiple years using revised protocols and more rigid quality control procedures.

Acknowledgements The authors would like to thank the readers from multiple institutes for the time and effort they put into reading the otoliths for this study—Jack Fife, Blanche Jackson, Colin MacDougall, Lisa Pinkham, Sarah Pierce, and Mike Power. We also thank Derek Knox for extracting the ageing material from the archives and for coordinating the exchanges. Jamie Joudrey provided the technical expertise in preparing the otoliths for radiocarbon assay.

References

Andrews AH, Burton EJ, Kerr LA, Cailliet GM, Coale KH, Lundstrom CC, Brown TA (2005) Bomb radiocarbon and lead-radium disequilibria in otoliths of bocaccio rockfish (*Sebastes paucispinis*): a determination of age and longevity for a difficult-to-age fish. Mar Freshw Res 56:517–528

Andrews AH, Kerr LA, Cailliet GM, Brown TA, Lundstrom CC, Stanley RD (2007) Age validation of canary rockfish (*Sebastes pinniger*) using two independent otolith techniques: lead-radium and bomb radiocarbon dating. Mar Freshw Res 58:531–541

Appelberg M, Formigo N, Geffen AJ, Hammer C, McCurdy W, Modin J, Moksness E, Mosegaard H, Morales-Nin B, Troadec H, Wright P (2005) A cooperative effort to exchange age reading experience and protocols between European fish institutes. Fish Res 76:167–173

Beamish RJ, McFarlane GA (1983) The forgotten requirement for age validation in fisheries biology. Trans Am Fish Soc 112:735–743

Beamish RJ, McFarlane GA (1995) A discussion of the importance of aging errors, and an application to walleye pollock: the world's largest fishery. In: Secor DH, Dean JM, Campana SE (eds) Recent developments in fish otolith research. University of South Carolina Press, Columbia, pp 545–565

Bertignac M, de Pontual H (2007) Consequences of bias in age estimation on assessment of the northern stock of European hake (*Merluccius merluccius*) and on management advice. ICES J Mar Sci 64:981–988

Bradford MJ (1991) Effects of ageing error on recruitment time series estimated from sequential population analysis. Can J Fish Aquat Sci 48(4):555–558

Bruch RM, Campana SE, Davis-Foust SL, Hansen MJ, Janssen J (2009) Lake sturgeon age validation using bomb radiocarbon and known-age fish. Trans Am Fish Soc 138:361–372

Campana SE (1984) Comparison of age determination methods for the starry flounder. Trans Amer Fish Soc 113(3):365–369

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 (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

Campana SE, Jones CM (1998) Radiocarbon from nuclear testing applied to age validation of black drum, *Pogonias cromis*. Fish Bull 96(2):185–192

Campana SE, Casselman JM, 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 65:733–743

Cleary L, Hunt JJ, Moores J, Tremblay D (1982) herring ageing workshop St. John's, Newfoundland March- 1982. Canadian Atlantic Fisheries Scientific Advisory Committee Research Document 82/41: 10p

Dery LM (2005) Atlantic herring, *Clupea harengus*. http://www.nefsc.noaa.gov/fbi/age-man/her/hertext.htm. 1–3

Dery L, Chenoweth J (1979) Recent problems in ageing sea herring from the Gulf of Maine. Reference Document 79–38, Woods Hole Laboratory. Northeast Fisheries Science Center, National Marine Fisheries Service, Woods Hole, MA 02543. 7p

Dunn A, Francis RICC, Doonan IJ (2002) Comparison of the Chapman-Robson and regression estimators of Z from catch-curve data when non-sampling stochastic error is present. Fish Res 59(1–2):149–159

Dwyer KS, Walsh SJ, Campana SE (2003) Age determination, validation and growth of Grand Bank yellowtail flounder (*Limanda ferruginea*). ICES J Mar Sci 60:1123–1138

Eklund J, Parmanne R, Aneer G (2000) between-reader variation in herring otoliths and effects on estimating population parameters. Fish Res 46(1–3):147–154

Francis MP, Campana SE, 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

Haas RE, Recksiek CW (1995) Age verification of winter flounder in Narragansett Bay. Trans Amer Fish Soc 124 (1):103–111

Hamel OS, Piner KR, Wallace JR (2008) A robust deterministic model describing the bomb radiocarbon signal for use in fish age validation. Trans Am Fish Soc 137:852–859

Hendrickson LC, Hart DR (2006) An aged based cohort model for estimating the spawning mortality of semelparous cephalopods with an application to per-recruit calculations for the northern shortfin squid, *Illex illecebrosus*. Fish Res 78(1):4–13

Horn PL (2002) Age and growth of Patagonian toothfish (*Dissostichus eleginoides*) and Antarctic toofish (D. *mawsini*) in waters from the New Zealand subantarctic to the Ross Sea, Antarctica. Fish Res 56(3):275–287

Hunt JJ, Parsons LS, Watson JE, Winters GH (1973) Report of herring ageing workshop. St. Andews, N.B., 11–13 December 1972. ICNAF Research Document 73/2:10p

Kalish JM (1993) Application off the bomb radiocarbon chronometer to the validation of redfish *Centroberyx affinis* age. Can J Fish Aquat Sci 52(7):1399–1405

Kalish JM (1995) Radiocarbon and fish biology. In: Secor DH, Dean JM, Campana SE (eds) Recent developments in fish otolith research. University of South Carolina Press, Columbia, pp 637–653

Kastelle CR, Kimura DK, Goetz BJ (2008) Bomb radiocarbon age validation of Pacific Ocean perch (*Sebastes alutus*) using new statistical methods. Can J Fish Aquat Sci 65:1101–1112

Kimura DK (1990) Approaches to age-structured separable sequential population analysis. Can J Fish Aquat Sci 47 (12):2364–2374

Kimura DK, Kastelle CR, Goetz BJ, Gburski CM, Buslov AV (2006) Corroborating the ages of walleye pollock (*Theragra chalcogramma*). Mar Freshw Res 57:323–332

Kneebone J, Natanson LJ, Andrews AH, Hunt HW (2008) Using bomb radiocarbon analyses to validate age and growth estimates for the tiger shark, Galeocerdo cuvier, in the western North Atlantic. Mar Biol 154(3):423–434

Libby DA, Burnett JM, Melvin GD (2006) Proceedings of the atlantic herring otolith age estimation workshop, 10–11 January 2006, West Boothbay Harbour, Maine. Transboundary Resource Assessment Committee Working Paper 2006/04. 18p

Maceina MJ, Boxrucker J, Buckmeier DL, Gangl RS, Lucchesi DO, Isermann DA, Jackson JR, Martinez PJ (2007) Current status and review of freshwater fish aging procedures used by state and provincial fisheries agencies with recommendations for future directions. Fisheries 32:329–340

McBride RS, Hendricks ML, Olney JE (2005) Testing the validity of Cating's (1953) method for age determination of American shad using scales. Fisheries 30(10):10–18

Melvin GD, Power MJ (2006) Expert opinion on the impact of age reading inconsistencies on the 2006 TAC Advice for Southwest NS / Bay of Fundy (SWNS/BoF) herring. Expert Opinion 06/06. 6p

Melvin GD, Power MJ (2007) Ageing inconsistencies and sensitivity analysis for 4WX herring. Can Sci Advis Secretariat Res Document 2007(54):1–35

Morison AK, Coutin PC, Robertson SG (1998a) Age determination of black bream, *Acanthopagrus butcheri* (Sparidae), from the Gippsland Lakes of southeastern Australia indicates slow growth and episodic recruitment. Mar Freshwater Res 49:491–498

Morison AK, Robertson SG, Smith DG (1998b) An integrated system for production fish aging: image analysis and quality assurance. N Am J Fish Manage 18:587–598

Neilson JD, Campana SE (2008) A validated description of age and growth of western Atlantic bluefin tuna (*Thunnus thynnus*). Can J Fish Aquat Sci 65(8):1523–1527

Overholtz WJ, Jacobson LJ, Melvin GD, Cieri M, Power MJ, Libby D, Clark, K (2004) Stock assessment of the Gulf of Maine-Georges Bank herring stock complex, 1959–2002. Northeast Fisheries Science Center, Reference Document 04–06. 300p

Parson LS, Winters GH (1972) ICNAF herring otolith Exchange 1071–72. Annu Meet int Comm Northw Alt Fish, Research Document 92, Serial No. 2816: 23p

Power MJ, Clark KJ, Fife FJ, Knox D, Melvin GD, Stephenson RL (2007) 2007 Evaluation of 4WX herring. Canadian Science Advisory Secretariat Research Document 2007/040. 79p

Power GR, King PA, Kelly CJ, McGrath D, Mullins E, Gullaksen O (2006) Precision and bias in the age determination of blue whiting, *Micromesistius poutas-*

sou, within and between age-readers. Fish Res 80(1–2):312–321

Reeves SA (2003) A simulation study of the impact of reading age errors for stock assessment and advice. ICES J Mar Sci 60(2):314–328

Restrepo VR, Power JE (1991) A comparison of three methods for handling the “plus” group in virtual population analysis in the presence of ageing errors. Collect Volume Scientific Papers Int Comm Conserv Atlantic Tunas 35(2):346–354

Richards LJ, Schnute JT, Kronlund AR, Beamish RJ (1992) Statistical models for the analysis of ageing error. Can J Fish Aquat Sci 49(9):1801–1815

Stewart REA, Campana SE, Jones CM, Stewart BE (2006) Bomb radiocarbon dating calibrates beluga (*Delphinapterus leucas*) age estimates. Can J Zool 84:1840–1852

Stuiver M, Polach HA (1977) Reporting of C-14 data. Radiocarbon 19:355–363

Sutherland SJ, Munroe NJ, Silva V, Pregracke SE, Burnett JM (2006) Accuracy and precision exercises associated with 2005 TRAC production aging. Northeast Fish Sci Cent Ref Document 06–28:8p

Tibbo SN (1970) ICNAF Scale and Otolith exchange program 1969–1970. Annu meet Int Comm Northw Atlant Fish 1970. Research Document No55, Serial No 2389: 23p

Treble MA, Campana SE, Wastle RJ, Jones CM, Boje J (2008) Growth analysis and age validation of a deepwater Arctic fish, the Greenland halibut (*Reinhardtius hippoglossoides*). Can J Fish Aquat Sci 65(6):1047–1059

Tyler AV, Beamish RJ, McFarlane GA (1989) Implications of age determination errors to yield estimates. Can Spec Publ Fish Aquat Sci 108:27–35

Watson JE (1965) A technique for mounting and storing herring otoliths. Trans Amer Fish Soc 94(3):267–268

Yule DL, Stockwell JD, Black JA, Cullis KI, Cholwek GA, Myers JT (2008) How systematic age underestimates can impede understanding of fish population dynamics: lessons learned from Lake Superior cisco stock. Trans Amer Fish Soc 137(2):481–495