



REVIEW PAPER

Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods

S. E. CAMPANA

Marine Fish Division, Bedford Institute of Oceanography, P.O. Box 1006, Dartmouth, Nova Scotia, Canada B2Y 4A2

(Received 6 March 2001, Accepted 21 May 2001)

Many calcified structures produce periodic growth increments useful for age determination at the annual or daily scale. However, age determination is invariably accompanied by various sources of error, some of which can have a serious effect on age-structured calculations. This review highlights the best available methods for insuring ageing accuracy and quantifying ageing precision, whether in support of large-scale production ageing or a small-scale research project. Included in this review is a critical overview of methods used to initiate and pursue an accurate and controlled ageing program, including (but not limited to) validation of an ageing method. The distinction between validation of absolute age and increment periodicity is emphasized, as is the importance of determining the age of first increment formation. Based on an analysis of 372 papers reporting age validation since 1983, considerable progress has been made in age validation efforts in recent years. Nevertheless, several of the age validation methods which have been used routinely are of dubious value, particularly marginal increment analysis. The two major measures of precision, average percent error and coefficient of variation, are shown to be functionally equivalent, and a conversion factor relating the two is presented. Through use of quality control monitoring, ageing errors are readily detected and quantified; reference collections are the key to both quality control and reduction of costs. Although some level of random ageing error is unavoidable, such error can often be corrected after the fact using statistical ('digital sharpening') methods.

Key words: age determination; otolith; accuracy; precision; quality; validation.

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 that of growth rate, or as complex as that of virtual population analysis, all require age data, since any rate calculation requires an age or elapsed time term. Radiochemical decay rates (Bennett *et al.*, 1982), lipofuscin accumulation rates (Hammer & Braum, 1988) and amino acid racemization rates (Goodfriend, 1992) are sometimes used to infer the age of a structure or organism, but in most cases, periodic growth increments are counted to estimate the age. Tree rings are the archetypal ageing structure, and have been used not only to determine age and date of formation, but through cross correlation with other trees, have been used to develop biochronologies

Tel.: +1902 426 3233; fax: +1902 426 9710; email: campanas@mar.dfo-mpo.gc.ca

extending over thousands of years (Kuniholm *et al.*, 1996). Annual varves in ice cores, sediments and stalagmites have been used to similar advantage (Baker *et al.*, 1993; Petterson *et al.*, 1999; Rittenour *et al.*, 2000). In the animal kingdom, annual or daily growth increments are used to estimate age and reconstruct growth rate in organisms and structures as diverse as bivalve shells (Lutz & Rhoads, 1980), coral skeletons (Dodge & Thomson, 1974), polychaete jaws (Olive, 1980), squid statoliths (Arkhipkin, 1997), cricket exoskeletons (Zuk, 1987), jellyfish statoliths (Ueno *et al.*, 1995), mammalian teeth (Goren *et al.*, 1987), brittlestar skeletons (Gage, 1990) and tortoise scutes (Germano, 1998). Where the growth increments have formed in calcified structures, environmental reconstruction based on incorporated trace elements and isotopes is also possible (Chivas *et al.*, 1985; Holmden *et al.*, 1997).

Several calcified structures produce periodic growth increments useful for age determination in fish. Scales (Robillard & Marsden, 1996), vertebrae (Brown & Gruber, 1988), fin rays (Cass & Beamish, 1983), cleithra (Casselman, 1990) and opercula (Baker & Timmons, 1991) have all been used to determine annual age, although it is the otolith which is applied over the broadest age range in many species (Secor *et al.*, 1995a). Campana & Thorrold (2001) estimated that well over 1 million fish were aged worldwide in 1999, most of those using scales and otoliths. Such efforts dwarf those routinely applied to non-fish species, and highlight the importance attributed to age-structured information in fisheries science.

Age determinations in fish can occur at one of two scales. Annual ageing is often used in support of harvest calculations and population studies, and can be based on any bony structure in the fish, although scales and otoliths are the structures most frequently used (Casselman, 1987). In contrast, daily ageing based on the otolith microstructure tends to be targeted more at recruitment questions and studies of young fish (Pannella, 1971; Campana & Neilson, 1985). Despite the difference in time scale, application and mode of formation, both annual and daily age data are governed by similar rules of analysis, and are susceptible to similar sources of error.

If growth increments in fish formed with the same consistency and clarity as those in temperate trees, and if the basis for fish growth was as clearly understood, population dynamics studies of fishes would be far more accurate than is now the case. Unfortunately, the process of estimating fish age incorporates two major sources of error: (a) a process error associated with the structure being examined; not all bony structures in fish form a complete growth sequence throughout the lifetime of the animal, nor do all axes within a given structure show a complete growth record (Beamish, 1979). This type of error is usually biased towards under- or over-ageing; and (b) error due to the element of subjectivity required of all age estimations. This subjectivity originates with the preparation and interpretation of the periodic features in the calcified structures, which can vary markedly among age readers and laboratories (Boehlert, 1985; Campana & Moksness, 1991). Interpretation error can be either biased or random. In combination, process and interpretation error can result in age estimates that differ by as much as a factor of three among investigators (Parrish, 1958; Campana *et al.*, 1990; Nedreaas, 1990; Donald *et al.*, 1992). Given the presence of such errors, the use of the term 'age determination' rather than 'age

estimation' would appear to be a bit of a misnomer. Nevertheless, the former term is in broad use around the world, and we will continue to use it here for the sake of familiarity.

The prevalence and impact of inaccurate age determinations on the accuracy of population dynamics studies cannot be overstated (Lai & Gunderson, 1987; Rivard & Foy, 1987; Tyler *et al.*, 1989; Bradford, 1991; Richards *et al.*, 1992; Morison *et al.*, 1998a). There are many instances in which ageing error has contributed to the serious overexploitation of a population or species. The problem is often one of age underestimation (rather than overestimation), resulting in overly optimistic estimates of growth and mortality rate. Examples include the orange roughy (*Hoplostethus atlanticus* Collett) off New Zealand that was fished intensively on the basis of a presumed longevity of 20–30 years (van den Broek, 1983). It is now suspected of living to over 100 years with an extremely slow growth rate (Smith *et al.*, 1995), but has already been fished almost to the point of population collapse. Similar problems plagued the *Sebastes* spp fisheries off eastern and western Canada, which are only now known to reach ages of over 75 years (Chilton & Beamish, 1982; Campana *et al.*, 1990), and thus less capable of supporting an intensive fishery. Ageing errors may also have contributed to errors in the population assessment of walleye pollock (*Theragra chalcogramma* Pallas) in the central Bering Sea, whose catches subsequently declined from 1 400 000 tons to 10 000 tons in less than a decade (Beamish & McFarlane, 1995). While the above-cited disasters are among the most visible examples of ageing inaccuracies, there are literally dozens of others cited in the literature which have resulted in serious scientific error (Summerfelt & Hall, 1987; Secor *et al.*, 1995a).

A number of authors have outlined methods through which ageing accuracy and/or objectivity can be improved, both at the daily (Brothers, 1979; Campana & Neilson, 1985; Geffen, 1987; Baillon, 1992) and yearly level (Blacker, 1974; Boehlert, 1985; Casselman, 1987; Cailliet, 1990). The past decade in particular has seen significant improvements in age determination protocols. At least some of these improvements can be attributed to Beamish & McFarlane's (1983) plea for age validation, in which they noted that only 66% of 500 publications reporting fish age estimates even attempted to corroborate the accuracy of their ages. A mere 3.4% were successful in doing so over the entire age range of the fish. The majority of published studies apparently assumed ageing accuracy, despite the fact that there was little basis for such an assumption.

Ageing error can be of two forms: error that affects accuracy, or the closeness of the age estimate to the true value, and error that affects precision, or the reproducibility of repeated measurements on a given structure (Kalish *et al.*, 1995). The two forms of error are not necessarily linked. For example, consistent underageing of a sample by one year can yield the same measure of precision as a sample that is, on average, aged accurately. In practice, the accuracy of a particular ageing methodology may be known ('age validation'), but the accuracy of a particular set of age estimates is seldom known. For these 'real world' samples, often consisting of large numbers of age determinations carried out at regular intervals [the 'production ageing' of Morison *et al.* (1998b)], relative accuracy may be just as important as absolute accuracy. For this reason, quality control monitoring is an important

component of any large-scale ageing program (Campana *et al.*, 1995; Morison *et al.*, 1998b).

The objective of this review is to highlight the best available methods for quantifying ageing accuracy and precision, whether in support of large-scale production ageing or a small-scale research program. Included in this review is a critical overview of methods used to initiate an accurate and controlled ageing program, including (but not limited to) validation of an ageing method. The overview will not consider the strategy or protocol for collecting age data; this topic has been well covered elsewhere (Chilton & Beamish, 1982; Morison *et al.*, 1998b). Rather, the focus will be on a series of protocols for quality control, primarily involving reference collections, so that any errors in ageing are quickly detected and corrected. The paper will then conclude with some statistical approaches for removing ageing error, and thus improving the quality of existing data.

ACCURACY AND AGE VALIDATION

The term ‘age validation’ has been used misleadingly in many past papers. Although the absolute age of the fish is the goal of validation studies, seldom is the age of the fish itself ever confirmed. Rather, it is the frequency of formation of a typical growth increment which is validated. The distinction between validating the periodicity of growth increment formation and absolute age is important. Beamish & McFarlane (1983) equated the validation of annulus periodicity with age validation, but then went on to state that all age groups must be validated before ageing accuracy can be accepted. If implemented rigorously, validation of annulus formation in each and every age group would be equivalent to validation of absolute age. However, such rigour has seldom (ever?) been displayed. In a recent glossary of otolith terminology, Kalish *et al.* (1995) were careful to note that age validation refers to validation of the method rather than the age, and that determining increment periodicity is only one part of the method. Nevertheless, the vast majority of published works equate confirmation of increment periodicity with age validation. Indeed, of 372 papers reporting age validation since the year of Beamish & McFarlane’s (1983) paper, only 15% actually validated the absolute age of wild fish. More than 50% validated growth increment periodicity for only a single group of ages, leaving increment periodicity unexamined for the most problematic groups: the oldest and/or youngest age groups. Yet it is the youngest and oldest fish which are often the most difficult to age accurately, and are most influential in estimates of growth, mortality or longevity.

Validation of an absolute age is equivalent to determining the accuracy of an age estimate. Determining the frequency of formation of a growth increment for a sample of fish is a necessary, but insufficient, step towards the verification of that age estimate. To illustrate this insufficiency, consider the following examples. Steffensen (1980) used otolith microstructure examination to infer the age of juvenile cod (*Gadus morhua* L.). Daily growth increment formation had already been validated in cod, so the frequency of increment formation was not in question. However, Steffensen did not confirm the age of formation of the first visible increment, and because of methodological problems, failed to

observe the first 90 increments. The result was a mean age which was about 50% of the actual age, despite the fact that he used a 'validated' method. In a second example, Pratt & Casey (1983) used various methods to infer growth increment periodicity on the vertebrae of mako sharks (*Isurus oxyrinchus* Rafinesque). Data were limited, but were consistent with the view that two increments formed each year in the vertebrae of the youngest sharks. Their subsequent examinations of the remaining mako vertebrae were thus based on the presumption of biannual increment formation, resulting in rapid apparent growth and low longevity for the oldest sharks, despite the fact that the validation was limited to the youngest age groups. We now know that the interpretation of vertebrae in young sharks is often problematic, and unlikely to be representative of subsequent growth (Natanson *et al.*, 2001). Yet the approach they used was considered (at the time) to have been validated.

Absolute age should be the preferred goal of any age validation study. Where this is not possible (and it often is not), two steps are recommended:

- (1) Determine the age of first increment formation. In many cases, this will require knowledge of the early life history of the fish, and will seldom be possible with the same experiment used to determine the frequency of increment periodicity. Even absolute age estimates are unlikely to provide sufficient precision to unequivocally identify the first annual (daily) increment.
- (2) Verify increment periodicity across the entire age range of interest. Growth increments of immature fish seldom resemble those of mature fish, as is evident in redfish (*Sebastes marinus* L.) otoliths (Fig. 1). Therefore a validation experiment confirming annulus formation in mature fish is unlikely to be applicable to immature fish, and vice versa. However, it is unrealistic, and probably unnecessary, to validate increment periodicity in every age group. At a minimum, validation of increment periodicity in both the youngest and oldest age groups is recommended. Of course, methods which estimate absolute age also verify increment periodicity across the entire age range.

Ageing accuracy does not necessarily result from the use of a fully validated ageing method. A validated ageing method confirms the frequency of formation of growth increments in a given structure, and confirms that they have been interpreted correctly by the age readers in the study. However, there is no guarantee that the same interpretation would be reached by other age readers, even when viewing the same structures. In other words, an age validation study deals with process error, but not with interpretation error. This interpretation error can be substantial. For example, six experienced ageing laboratories independently examined a set of prepared haddock (*Melanogrammus aeglefinus* L.) otoliths to determine age. Annulus formation and absolute age determination in this species has been validated (Campana, 1997), indicating that accurate age determinations are possible for even the oldest fish. Yet one of the laboratories consistently underestimated the ages of most of the mature fish by about 50% compared to the other laboratories (Campana, 1995). The ageing method was validated, all age readers were looking at identical preparations, yet the interpretations of one of the laboratories was vastly different. In another

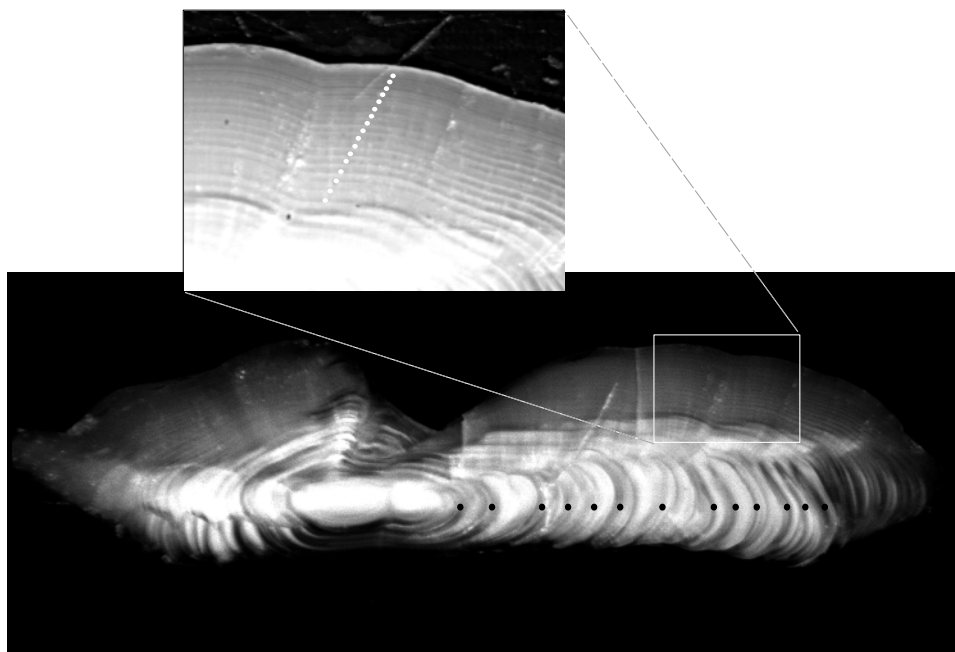


FIG. 1. Transverse section of a redfish (*Sebastes mentella*) otolith, revealing the annuli. Annuli formed before sexual maturation (●) look very different than those formed after sexual maturation (○; insert). As a result, an age validation experiment targeting immature fish is unlikely to help in the interpretation of annuli of mature fish, and vice versa.

example, [Gauldie *et al.* \(1993\)](#) published a paper reporting that annuli were not formed in the otoliths of two known-age fish reared in a public aquarium for more than 16 years. Yet when the otolith photographs included in the paper were shown to a variety of experienced age readers, most were able to determine the exact age of the fish based only on the photographs. Once again, this illustrates that interpretive error can remain as a significant source of error in an age validation study, other than by those who carried out the study. In other words, citation of another published study in support of one's own ages does not necessarily imply that your ages are accurate.

How then does one go about validating not just the ageing structure, but the correct interpretation thereof? Chemists have long used certified reference materials (CRMs) to confirm that independent laboratories were providing comparable assays, despite any differences in methodology or instrumentation ([Beauchemin *et al.*, 1987](#)). Until recently, it was difficult for fisheries scientists to ensure similar comparability, other than through exchange of photographs or the actual structures used in the age validation study. However, with the popularity of image analysis systems for acquiring images, and the advent of the World Wide Web, image exchange has become both routine and fast. There is now little reason why images of the structures used in a validation study could not be posted on a Web site at the time of publication for others to examine. Thus all interested parties could examine validated material to insure that their own interpretations were accurate. Through such a mechanism, an age validation study would serve not only to validate growth increment periodicity, but to

validate the interpretation of any and all age readers. The Web-posted images would then become the equivalent of an image CRM, and would do much to improve ageing accuracy around the world.

AGE VALIDATION METHODS

A variety of methods exist through which age interpretations can be validated (Table I). 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. Several reviews of age validation techniques suitable for annual (Blacker, 1974; Bagenal & Tesch, 1978; Casselman, 1983, 1987; Cailliet *et al.*, 1986; Beamish & McFarlane, 1987; Baillon, 1992; Campana, 1999) and daily ages (Brothers, 1979; Campana & Neilson, 1985; Jones, 1986; Geffen, 1987, 1992) have been published. Below is presented a critical appraisal of the various approaches, ordered (subjectively) by scientific value, with suggestions for enhancing scientific rigour.

RELEASE OF KNOWN AGE AND MARKED FISH

Release of known age and marked fish into the wild is probably the most rigorous of the age validation methods for many species, since the absolute age of the recaptured fish is known without error. Since the released fish are generally less than 1 year old, recaptured fish will have spent the majority of their lives in natural surroundings. Fish can be marked either externally, as in the case of salmon with coded wire tags (Quinn *et al.*, 1991), or immersion mass-marked using temperature fluctuations (Volk *et al.*, 1999) or chemicals (Campana, 1999) so as to leave a permanent mark on the bony structures used for ageing. This approach is not well suited to long-lived species, since recapture rates of old fish tend to be minimal. Nor can this method be used on species which cannot be reared in captivity prior to release. Nevertheless, this method has been used with success to confirm absolute age and growth increment formation at both the daily (Tsukamoto & Kajihara, 1987; Secor *et al.*, 1995b) and the yearly scale (Fitzgerald *et al.*, 1997; Svedang *et al.*, 1998).

There are two variations on this method which make it more widely available at the expense of relatively minor assumptions. The first variation involves scale removal at the time of tagging and release of wild fish. Where tagging has been restricted to relatively young fish, and where scale annuli have been found to be reliable indicators of age at that young age, the removed scale can be used to estimate the age at tagging, and subsequently be added to the time at liberty to estimate the absolute age of the fish. Where the age at tagging is short compared to the time at liberty, the advantage of this approach is that the wild tagged fish effectively become known age at release and thus need not be reared in captivity (e.g. Matlock *et al.*, 1993). A second variation on this theme involves the tagging of young fish where age can reasonably be approximated by size. This approach was used by Lee & Prince (1995) in their study of bluefin tuna (*Thunnus thynnus* L.), whereby tuna estimated to be 1–3 years old at the time of tagging were subsequently recaptured up to 15 years later. Although there was a ± 1 year

TABLE I. Features, advantages and disadvantages of methods used to confirm or support the accuracy of age interpretations. Methods are listed in descending order of scientific value. Growth structure refers to either annulus (A) or daily growth increment (D), depending on application

Method	Age validation							
	Annual/ daily	Applicable age range	Advantages	Limitations	Precision	Sample size required	Time required	Cost
Release of known-age and marked fish into wild	AD	all	<ul style="list-style-type: none"> validates both absolute age and periodicity of growth structures well suited to fishes with a longevity of < 10 years 	<ul style="list-style-type: none"> requires source of known-age and marked fish number of recaptures of old fish can be low or non-existent 	± 0 year	>1	>1–10 years	minimal if fish source is available
Bomb radiocarbon	A	all	<ul style="list-style-type: none"> validates both absolute age and periodicity of growth structures well suited to long-lived fishes does not require recently-collected samples 	<ul style="list-style-type: none"> at least some of fish in sample must have been hatched before 1965 	$\pm 1-3$ years	20–30	<1 year	\$700–\$1000 per otolith
Mark-recapture of chemically-tagged wild fish	AD	all	<ul style="list-style-type: none"> validates periodicity of post-tagging growth structures in fish of any age 	<ul style="list-style-type: none"> number of recaptures of fish at liberty more than 1 year can be low or non-existent identification of a single post-mark annulus can be problematic 	± 1 year	>1	>1–10 years	minimal excluding cost of tagging cruise

TABLE I. *Continued*

Age validation								
Method	Annual/ daily	Applicable age range	Advantages	Limitations	Precision	Sample size required	Time required	Cost
Radiochemical dating	A	5+ years	<ul style="list-style-type: none"> • validates absolute age • can be applied to any recently-collected samples • well suited to long-lived fishes 	<ul style="list-style-type: none"> • can only distinguish between widely divergent age estimates 	$\pm 25\text{--}50\%$	10–50	< 1 year	~\$1000 per age category
Progression of discrete length modes sampled for age structures	AD	0–5 years	<ul style="list-style-type: none"> • well suited for validating the first 1–2 age classes 	<ul style="list-style-type: none"> • length mode must not overlap that of adjacent mode • assumes no size-selective immigration or emigration into the sampling area 	± 0 year	> 100	1 year	minimal other than fish collection
Capture of wild fish with natural, date-specific markers	AD	all	<ul style="list-style-type: none"> • validates periodicity of growth increments and sometimes absolute age 	<ul style="list-style-type: none"> • natural, date-specific markers are very rare 	± 0 year	> 1	> 1 year	minimal
Marginal increment analysis	A	all	<ul style="list-style-type: none"> • validates periodicity of growth increments 	<ul style="list-style-type: none"> • only suited to fast-growing and/or young fish • requires samples from throughout year 	± 1 year	> 100	1 year	minimal other than fish collection
Captive rearing from hatch	AD	all	<ul style="list-style-type: none"> • validates both absolute age and periodicity of growth structures 	<ul style="list-style-type: none"> • otolith increments in reared fish seldom resemble those of wild fish 	± 0 year	> 1	1–10 years	minimal
Captive rearing of chemically-tagged fish	AD	all	<ul style="list-style-type: none"> • validates periodicity of growth increments • particularly easy for daily increments 	<ul style="list-style-type: none"> • otolith increments in reared fish seldom resemble those of wild fish 	± 0 year	> 1	1–10 years	minimal

TABLE I. *Continued*

Corroboration of age interpretation								
Method	Annual/daily	Applicable age range	Advantages	Limitations	Precision	Sample size required	Time required	Cost
Tag-recapture analysis	AD	all	<ul style="list-style-type: none"> provides excellent growth comparison for well-sampled age classes 	<ul style="list-style-type: none"> fish at liberty the longest provide the most useful data, but are seldom recaptured 	$\pm 0-10$ years	> 1	1-10 years	minimal excluding cost of tagging cruise
Length frequency analysis	AD	< 7 years	<ul style="list-style-type: none"> takes advantage of data most often available (length) can provide good growth rate estimates for first few age classes 	<ul style="list-style-type: none"> assumes one spawning period per year useful only for fast-growing fish where age-specific length modes are discernible 	$\pm 0-1$ year	> 100	< 1 year	minimal other than fish collection
Progression of strong year-classes	A	1-20 years	<ul style="list-style-type: none"> provides rapid, inexpensive but qualitative view of ageing accuracy 	<ul style="list-style-type: none"> strong (or weak) year-classes eventually disappear from catch at age through normal ageing imprecision 	± 3 year	> 1000	< 1 year	minimal if catch at age available
Numerical integration of daily growth increment widths	A	1-7 years	<ul style="list-style-type: none"> provides estimate of absolute age and growth rate in the absence of any other information where annuli are visible, can be used to validate their periodicity 	<ul style="list-style-type: none"> difficult to satisfy assumptions that daily increment sequence is uninterrupted and that unseen daily increments are similar in width to observed increments 	$\pm 0-1$ years	> 1	< 1 year	minimal

TABLE I. *Continued*

		Corroboration of age interpretation						
Method	Annual/ daily	Applicable age range	Advantages	Limitations	Precision	Sample size required	Time required	Cost
Daily increments between annuli	A	1	<ul style="list-style-type: none"> valuable for identifying 1st annulus 	<ul style="list-style-type: none"> pre-supposes knowledge of dates of hatch and annulus formation difficult to satisfy assumption that cycles are induced only by the environment not suitable for slow-growing fish, since annuli are too narrow 	± 0 year	> 1	< 1 year	minimal
Elemental and isotopic cycles	A	1–15 years	<ul style="list-style-type: none"> can be correlated with environmental cycles to infer age 	<ul style="list-style-type: none"> assumes no immigration, emigration, recruitment or age-specific mortality 	$\pm 0-1$ year	> 1	< 1 year	\$50–\$500 per otolith
Interval between samples	AD	all	<ul style="list-style-type: none"> can be used to validate periodicity of growth increments 	<ul style="list-style-type: none"> assumes no immigration, emigration, recruitment or age-specific mortality 	$\pm 0-1$ year	> 100	1 + year	minimal other than fish collection

margin of error around the age estimate at the time of tagging, that margin was too small to change the conclusion that vertebral growth marks were formed annually after tagging.

BOMB RADIOCARBON

Bomb derived radiocarbon from nuclear testing provides one of the best age validation approaches available for long-lived fishes (Kalish, 1993, 1995a, b; Kalish *et al.*, 1996, 1997; Campana, 1997, 1999; Campana & Jones, 1998). The onset of nuclear testing in the late 1950s resulted in an abrupt increase in atmospheric ^{14}C , which was soon incorporated into corals, bivalves, fish and other organisms that were growing at the time. Thus the period is analogous to a large-scale chemical tagging experiment, wherein all otolith cores of fish hatched before 1958 contain relatively little ^{14}C and all those hatched after 1968 contain elevated levels. Fish born in the transition period contain intermediate levels. As a result, the interpretation of the ^{14}C chronology in a sample of otolith cores is relatively simple; the otolith-based ^{14}C chronology spanning the 1960s should match other published ^{14}C chronologies for the region (whether from otoliths or other calcified organisms) as long as the annular age assignments (=year-class) are correct. Any under-ageing would phase shift the otolith ^{14}C chronology towards more recent years, while over-ageing would phase shift it towards earlier years. Otolith contamination with material of more recent origin can only increase the $\Delta^{14}\text{C}$ value, not decrease it. Thus the otolith $\Delta^{14}\text{C}$ value sets a minimum age to the sample, and the years 1958–1965 become the most sensitive years for $\Delta^{14}\text{C}$ -based ageing. For fish born during this time period, bomb radiocarbon can be used to confirm the accuracy of more traditional ageing approaches with an accuracy of at least ± 1 –3 years; the discriminatory power of samples born before or after this period is more than an order of magnitude lower. Since the ^{14}C signal recorded in deepsea and freshwater environments is different from that of surface marine waters (deepsea=delayed; fresh water=advanced), reference ^{14}C chronologies appropriate to the environment experienced during the period of otolith core formation must be used (Kalish, 1995b; Campana & Jones, 1998). Clearly, this approach is not well suited to studies of short-lived (<5 years) species, in instances where the presumed hatch dates do not span the 1960s, or in environments where appropriate reference chronologies are not available. On the other hand, the low radioactive decay rate of ^{14}C implies that both archived and recent collections are appropriate for assay.

MARK-RECAPTURE OF CHEMICALLY-TAGGED FISH

Mark-recapture of chemically-tagged (OTC) wild fish is one of the best methods available for validating the periodicity of growth increment formation. The method is based on rapid incorporation of calcium-binding chemicals such as oxytetracycline, alizarin, calcein or strontium, applied at the time of tagging, into bones, scales, spines and otoliths (Campana, 1999). Application is through immersion, injection or feeding, although injection is the most practical method for tagging studies of wild fish (Geffen, 1982; Foreman, 1987; Francis *et al.*, 1992; Oliveira, 1996). The result is a permanent mark, visible under fluorescent light (except strontium), in the growth increment being formed at the time of

tagging. The number of growth increments formed distal to the chemical mark is then compared to the time at liberty after tagging. This approach has been used to validate annulus formation in a wide variety of structures and species, including sablefish otoliths (Beamish & Chilton, 1982), shark vertebrae (Brown & Gruber, 1988), pike cleithra (Casselmann, 1974), spiny dogfish spines (Beamish & McFarlane, 1985), and coral reef fish otoliths (Fowler, 1990). The approach has also been used successfully at the microstructural level, validating daily increment formation in a variety of tuna species (Wild & Foreman, 1980; Laurs *et al.*, 1985). A major advantage of this approach is that the growth increments being validated are formed while the fish is growing in a natural environment. Experiments in which fish are chemically-tagged and then reared in the laboratory or an outside enclosure (Campana & Neilson, 1982; Schmitt, 1984) are less optimal, although they are logistically easier to carry out. A disadvantage of the chemical tagging approach is that the number of increments formed after tagging is often low, resulting in a potentially large relative error if one of the increments (such as that at the growing edge) is misinterpreted. For example, misinterpretation of a single growth zone in a fish at liberty 2 years would result in a 50% error, whereas the same misinterpretation in a fish at liberty 10 years would only produce a 10% error. This effect was highlighted in a recent study in which long-term mark-recaptures detected problems with annulus identification that were not evident from short-term recaptures in the same study (Beamish & McFarlane, 2000). For this reason, fish tagged at a young age and recaptured at an old age provide the most robust validation results (Natanson *et al.*, 2001). Notwithstanding the caveat that this method only validates growth increment formation for the size/age of fish tagged, this is a powerful method, and one of the few readily applied to adult wild fishes.

RADIOCHEMICAL DATING

Radiochemical dating of otoliths is based on the radioactive decay of naturally occurring radioisotopes which are incorporated into the otolith during its growth. Once incorporated into the otolith, the radioisotopes decay into radioactive daughter products, which are themselves retained within the acellular crystalline structure. Since the half-lives of the parent and daughter isotopes are known (and fixed), the ratio between them is an index of elapsed time since incorporation of the parent isotope into the otolith. By restricting the assay to the extracted otolith core (as opposed to the whole otolith), objective, accurate estimates of absolute age are possible (Bennett *et al.*, 1982; Campana *et al.*, 1990, 1993; Fenton *et al.*, 1990, 1991; Smith *et al.*, 1991; Kestelle *et al.*, 1994; Milton *et al.*, 1995; Burton *et al.*, 1999; Campana, 1999). The isotopic concentrations requiring measurement are exceedingly low, resulting in assay precisions which are often less than optimal, although recent methodological changes have substantially improved precision (Andrews *et al.*, 1999). Current discriminatory power is on the order of 5 years for $^{210}\text{Pb} : ^{226}\text{Ra}$ and 1–2 years for $^{228}\text{Th} : ^{228}\text{Ra}$, over age ranges of 0–40 and 0–8 years, respectively. Therefore, this approach is best suited to long-lived species where the candidate age interpretations are widely divergent, such as in *Sebastes* or *Hoplostethus* (Campana *et al.*, 1990; Fenton *et al.*, 1991).

DISCRETE LENGTH MODES SAMPLED FOR AGE STRUCTURES

Progression of discrete length modes sampled for age structures has seldom been applied rigorously, but it is a reasonably robust approach for validating the interpretation of annuli in young fish. By monitoring the progression of discrete length modes across months within a year, it is relatively straight forward to determine if the modes correspond to age classes (Natanson *et al.*, 2001). In instances where the length modes are well separated, can be tracked throughout the year, are not confounded by size-selective mortality, migration or multiple recruitment pulses within a year, and the mode corresponding to the young-of-the-year can be unequivocally identified, absolute age is confirmed. Examination of the ageing structures sampled from those same modes can then be used to test the validity of the presumed annuli as age indicators. This was the basis of the approach by Hanchet & Uozumi (1996), who found good correspondence between the number of presumed annuli and the age of the first three well-defined length modes (where the age was confirmed by modal progression). This approach is not equivalent to that which is more commonly applied, in which discrete length modes observed in a single sample are each assumed to correspond to an age class (Shirvell, 1981; Morales-Nin, 1989). While such an approach provides corroboration for an age interpretation, there is no independent evidence that the length modes represent age classes; thus strictly speaking, an approach that does not track modal progression through the year does not validate either absolute age or annulus periodicity. In principle, sampled modal progression should also be applicable to daily age validation. In practice however, size-selective mortality and/or migration is often pronounced in young fishes, thus invalidating the assumption that a distinct cohort is being tracked (Meekan & Fortier, 1996).

NATURAL, DATE-SPECIFIC MARKERS

Capture of wild fish with natural, date-specific markers is an approach that has many of the same advantages and disadvantages of bomb radiocarbon dating, since it relies on a large-scale event that applies a dated mark to all fish in a population. In the specific (and rare) instances in which it can be applied, this method can be used to validate growth increment formation over a substantial portion of a fish's life history. For example, both Blacker (1974) and Rauck (1974) reported the presence of otolith annuli which appeared to be characteristic of specific year-classes, such as the characteristically narrow second year growth zone of 1 year-class of Bear Island cod. More recently, MacLellan & Saunders (1995) suggested that the El Niño-induced disruption of growth in one year-class of Pacific hake (*Merluccius productus* Ayres) could be used as a dated marker to validate the frequency of annulus formation in fish from this year-class as it grew older. In general however, such marks would seldom be expected to be unambiguously identifiable in individual fish, and in any event, would have to be monitored over a number of years to insure that the mark remained identifiable.

A related but different approach is to take advantage of physiologically-generated marks or checks on the ageing structure, such as the hatching, emergence or first feeding check of salmonids (Marshall & Parker, 1982). This can be a powerful validation method of either absolute age or increment periodicity, as long as the date of check formation can be determined through

independent observation, and as long as the identity of the check is unambiguous. In the salmonid example above, once an observer had noted the date of emergence of a specific fish from the gravel, the emergence check had become a dated mark on the otolith of that fish that could then be used to validate both the absolute age and the frequency of formation of the daily increments formed until the date of capture. This method is probably better suited to daily increment validation than to annulus validation, since hatch checks (Campana & Neilson, 1985) and settlement marks (Wilson & McCormick, 1997) are common in some groups of fishes. Nevertheless, analogous marks do exist in many older fish, such as the otolith transition zone associated with the onset of sexual maturity (Francis & Horn, 1997). In all cases however, a key requirement is the independent observation of the date of the physiological event, since without it, the check is associated with an age, but not a date of formation.

MARGINAL INCREMENT ANALYSIS

Marginal increment analysis (MIA) is the most commonly used, and the most likely to be abused, of the validation methods. The underlying premise as a method for validating increment periodicity is sound: if a growth increment is formed on a yearly (daily) cycle, the average state of completion of the outermost increment should display a yearly (daily) sinusoidal cycle when plotted against season (time of day) (e.g. Hyndes *et al.*, 1992; Fowler & Short, 1998; Morales-Nin *et al.*, 1998; Carlson *et al.*, 1999). The popularity of this method can be attributed to its modest sampling requirements and low cost. However, in many ways, this is one of the most difficult validation methods to carry out properly, due to the technical difficulties associated with viewing a partial increment affected by variable light refraction through an edge which becomes increasingly thin as the margin is approached, as well as light reflection off the curved surface of the edge. The absence of an objective means of interpreting the data further complicates the situation. In their review of annulus seasonality studies, Beckman & Wilson (1995) interpreted the results of 104 MIA studies, concluding that about 30% of the species from a given region formed annuli at times different than that of the other species. It is possible that annuli did not form in all of these species, or that the time of opaque zone formation varied widely among species. Indeed, Beckman & Wilson (1995) highlighted the current lack of understanding of the mechanisms underlying annulus formation. However, a more likely explanation is that the MIA technique itself was of low resolving power. Even more problematic are studies which attempt to validate daily increment formation with MIA, working near the resolution limit of light, and confounded by the presence of subdaily increments. Although daily MIA based on transmission electron microscopy (Zhang & Runham, 1992) or using otoliths with unusually broad increments (Jenkins & Davis, 1990) has some merit, MIA studies of daily increments are, in general, of questionable value.

Marginal increment analysis is sometimes differentiated from edge analysis, but when used as a validation method, has similar properties. The marginal increment is usually calculated as a proportional state of completion, ranging from near zero (an increment is just beginning to form) to one (a complete increment has formed) as well as all values in between. When plotted as a function of month or season, the mean marginal increment should describe

a sinusoidal cycle with a frequency of one year in true annuli (Lehodey & Grandperrin, 1996; Vilizzi & Walker, 1999). Edge analysis does not assign a state of completion to the marginal increment, but rather records its presence as either an opaque or translucent zone (van der Walt & Beckley, 1997; Labropoulou & Papaconstantinou, 2000). It is the change in relative frequency of each edge zone which is plotted across months or seasons, but as with MIA, the cycle frequency should equal one year in true annuli. In both MIA and edge analysis, a yearly cycle of formation can be difficult to distinguish from other frequencies, contributing to their poor performance as validation methods. Changes in the seasonal timing of the marginal increment with age or location undoubtedly contribute to the problem; significant and unexplained differences among years have also been noted (Pearson, 1996; Cappo *et al.*, 2000). Despite the problems inherent in their use for age validation, both MIA and edge analysis are well suited for determining the month or season of formation of the opaque or translucent zone once annulus formation has been validated through independent means (Pearson, 1996; Natanson *et al.*, 2001).

There are several reasons why MIA may provide misleading results. Prominent among these is the fact that the marginal increment is most easily discerned in young, fast-growing fish, a life history stage where the marginal increment may accurately confirm the formation of annual increments. The problem arises when the 'validation results' are later applied to older fish, contrary to the assumptions of all age validation methods. Many studies have reported age validation based on MIA of young fish, but noted that the same ageing structure and/or approach provided incorrect ages in older fish (Campana, 1984; Hyndes *et al.*, 1992; Lowerre-Barbieri *et al.*, 1994). More troublesome are the instances where age validation based on MIA of young fish later evolved to form the basis for routine ageing of the species across all age groups. For example, MIA of scales in young snapper (*Pagrus auratus* Bloch and Schneider) quickly evolved to become the basis for all scale ageing of the species in several countries; OTC mark-recapture results later showed that scale ages underestimated true age in older fish (Francis *et al.*, 1992). A nearly identical situation took place in the north-east Pacific, where all routine ageing of sablefish (*Anoplopoma fimbria* Pallas) by several countries was based on scales validated with MIA. It wasn't until otolith OTC mark-recapture studies were completed that it was realized that scale ages were underestimating the age of older fish by up to a factor of four (McFarlane & Beamish, 1995). Note however that MIA misuse is not restricted to scale ageing. Annuli in whole otoliths of redfish (*Sebastes* spp.) were validated using MIA, and subsequently became an accepted procedure of many organizations for ageing these long-lived fishes; subsequent validations have demonstrated that whole otoliths grossly underestimate age in older fish (Campana *et al.*, 1990). The conclusion is clear: when proper age validation studies are lacking, researchers will often seize upon any available studies which can corroborate their age interpretations. And since MIA is one of the few validation methods which is restricted to young, fast-growing fish, it is also the most likely to lead to serious ageing error when applied blind.

It is difficult to recommend the use of a technique where the data can be so subjectively interpreted. Nonetheless, the approach is valid if done with

sufficient rigour. Four aspects of a rigorous protocol appear to be important: (1) samples must be completely randomized before examination, with no indication to the examiner when the sample was collected; (2) a minimum of two complete cycles needs to be examined, in accordance with accepted methods for detecting cycles; (3) the results must be interpreted objectively, extending well beyond the 'looks like a cycle to me' interpretation that is so commonly used. It is difficult to recommend one statistical test that would apply in all circumstances, although a variety of useful options have been offered (Vilizzi & Walker, 1999; Cappelletti *et al.*, 2000). At a minimum however, there should be significant differences among some or all of the seasonal groups in each of the cycles examined; and (4) the MIA should be restricted to only a few age groups at a time, ideally only one. As noted by Hyndes *et al.* (1992) in a study of whole otoliths, examination of a sample which includes young, annulus-producing fish and older, non-annulus producing fish can easily result in a significant annual cycle for the sample as a whole, despite the fact that the older fish by themselves would not show such a cycle. In other words, the validation results should be considered to be age-specific.

CAPTIVE REARING

Captive rearing is generally discounted as a reliable means of validating annulus formation, but maintains some utility at the daily level. Laboratory environments are seldom able to mimic natural environments, due to their artificial photoperiods, temperature cycles, feeding schedules and limited space for diurnal vertical migrations. Since annulus formation is strongly influenced by the environment (Schramm, 1989; Beckman & Wilson, 1995), an artificial environment is likely to produce artificial annuli. Daily growth increments are much less affected by environmental conditions, due to the endocrine-driven endogenous rhythm which controls their formation (Campana & Neilson, 1985). While laboratory environments are well known for resulting in daily increments of altered appearance, the frequency of their formation is not generally an issue unless the rate of growth is unnaturally low. For this reason, laboratory experiments to confirm daily increment formation of known-age or chemically-marked fish are common (Geffen, 1992).

Mesoscosms, ocean pens and outside enclosures provide improved and more natural rearing environments for validation studies than do indoor locations. For otolith microstructure studies in particular, outdoor rearing can be expected to produce daily increments which are quasi-natural in appearance and frequency, although growth rates can be artificially high in hatchery operations (Campana & Neilson, 1982; Folkvord *et al.*, 1997). At the annual level, outdoor rearing can also be expected to produce more natural-looking growth structures, although it has not yet been determined if annuli produced under such conditions are equivalent to those of wild fish (Schramm, 1989).

FREQUENCY OF USE OF AGE VALIDATION METHODS

Each age validation method has advantages and disadvantages (Table I) which would be expected to influence the frequency of their use. Perhaps not

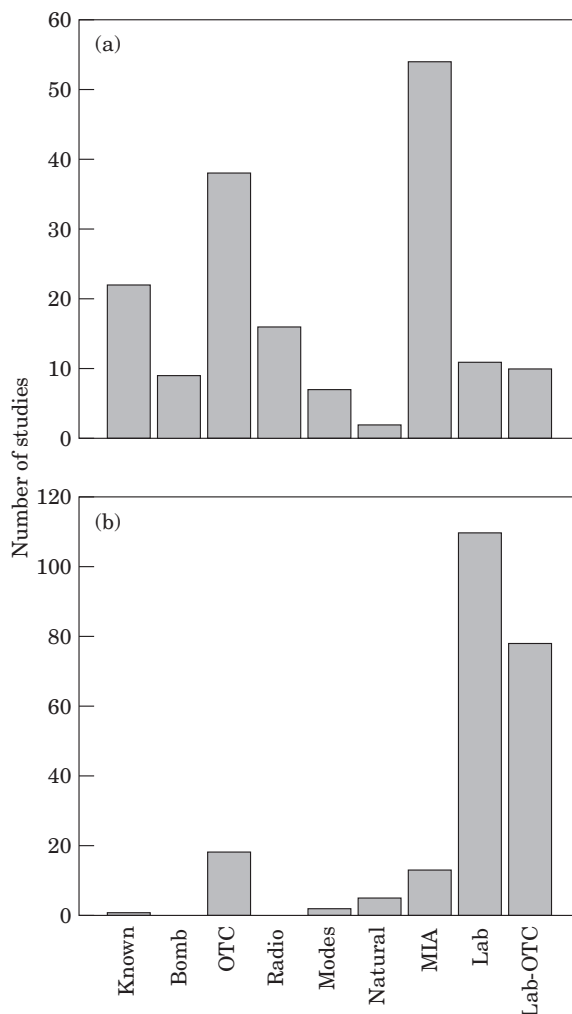


FIG. 2. Summary of papers reporting age validation after 1983, categorized by age validation method: (a) annual ages ($n=205$); (b) daily ages ($n=162$). Known, known age; Bomb, bomb radiocarbon; OTC, mark-recapture of chemically tagged wild fish; Radio, radiochemical dating; Modes, progression of length modes sampled for ages; Natural, natural, date-specific markers; MIA, marginal increment analysis; Lab, captive rearing from hatch; Lab-OTC, captive rearing after chemical marking.

surprisingly though, cost and opportunity appear to play major roles in the selection of a method. In a review of 372 papers published since the appearance of [Beamish & McFarlane's \(1983\)](#) plea for age validation, the majority of papers attempting to validate annuli used MIA, one of the least rigorous methods, to do so ([Fig. 2](#)). On a more encouraging note, more than 40% of the annulus validation studies used one of the three most rigorous methods, a substantial improvement over the situation prior to 1983.

Annulus validation studies were most focused on otoliths ($n=102$), with smaller numbers dealing with vertebrae ($n=26$), scales ($n=10$), fin spines/rays

($n=9$), bivalves ($n=7$), and others ($n=13$). However, validation methods were not equitably distributed among ageing structures. More than 44% of the otolith annulus studies used MIA for validation, while only 17% of the studies on other ageing structures did so. Scale validation studies in particular tended to use very rigorous techniques (known age, OTC mark-recapture), presumably because of their well documented failures in some species in the past (Beamish & McFarlane, 1987; Casselman, 1987).

Validation methods for daily growth increments require different approaches than do those for annuli, so the selection of method would also be expected to differ. More than 90% of the reported studies reared organisms in captivity to validate daily increment formation, either from hatch or after chemical marking (Fig. 2). As noted earlier, captive rearing is of greater scientific rigour in daily increment validation studies than in those on annuli, due to the lesser influence of the environment on frequency of formation (Campana & Neilson, 1985). Nevertheless, the relative scarcity of validation studies on wild fish was somewhat surprising. Otolith papers made up the bulk of the studies ($n=186$), but studies on daily increments in squid statoliths ($n=16$), bivalves ($n=2$) and scales ($n=1$) used similar approaches. MIA was seldom used (6%), and then primarily in otoliths.

With greater scientific quality, it is perhaps not surprising that some recent age validation results appeared to contradict earlier, less rigorous results. For example, Beamish & McFarlane (2000) reevaluated annulus formation on sablefish otoliths after extended periods at liberty after OTC tagging, and reported that annulus formation was not necessarily as clearcut as had been reported earlier (McFarlane & Beamish, 1995). The difference in interpretation was apparently due to the relatively short period after recapture in the earlier study, resulting in poor sensitivity. In a second study, Wild *et al.* (1995) reported that daily increment formation became inconsistent in large skipjack tuna (*Katsuwonus pelamis* L.), despite the results of an earlier validation study (Uchiyama & Struhsaker, 1981). The 1981 study was based on young, lab-reared tuna, while the 1995 report was based on a much more rigorous study of older, OTC-marked fish in the wild. Campana (1983) also reported non-daily increment formation in some starry flounder (*Platichthys stellatus* Pallas) otoliths, apparently contradicting an earlier report of daily increment validation (Campana & Neilson, 1982). In this case however, the difference in results was not due to a change in scientific rigour but to tests on starved fish. Nevertheless, these three sets of studies highlight the fact that not all age validation studies are created equal, and that increased scientific rigour will always produce more reliable results.

In summary, considerable progress appears to have been made since Beamish & McFarlane's (1983) paper in attempts to validate age interpretations. The increased frequency of studies to validate absolute annual age and use chemical mark-recapture are particularly encouraging. Nevertheless, additional work is clearly required. The heavy reliance and apparent abuse of MIA, particularly in the otolith world, is disturbing. With respect to daily increment studies, additional studies using known age and chemically marked fish in their natural environment appear to be needed.

CORROBORATION OF AGE INTERPRETATIONS

Methods for age corroboration are not equivalent to those for age validation, since corroboratory methods support or are correlated with a particular method of ageing, but are not directly or logically linked. As a result, it is entirely possible to have an age corroboration method which reinforces an incorrect age interpretation. Nevertheless, a well designed corroboratory study can provide valuable support for a proper age validation study in confirming the accuracy of an age estimate or method. The following briefly summarizes the major age corroboration methods that are currently available, some of which have previously been called (incorrectly) age validation methods. Advantages and disadvantages of each approach are also shown in [Table I](#).

TAG-RECAPTURE ANALYSIS

Tag-recapture analysis, along with length frequency analysis, is a member of a suite of methods which provides growth rate estimates which can be compared with those derived from annulus counts. The growth comparison is by inference, since none of the recaptured fish are of known age. Nonetheless, if sufficient tag returns are available, and particularly if the capture and release sizes were carefully measured, the resulting growth rate estimate is an important check on the accuracy of the age determination method. The traditional method of [Gulland & Holt \(1959\)](#) uses a graph of annualized growth rate after tagging plotted against average length between tagging and recapture to calculate the von Bertalanffy growth parameters, L_{∞} and K . This method has been widely used (e.g. [Thorson & Lacy, 1982](#); [Natanson *et al.*, 1999](#)), but assumes von Bertalanffy growth, no measurement error and no seasonal variability in growth rate. A more rigorous approach is the GROTAG analysis of [Francis \(1988\)](#), which uses maximum likelihood methods to estimate growth rate, growth variability and measurement error at two lengths. The approach properly differentiates between growth at length and growth at age, and produces reliable growth estimates, but at the cost of considerable recapture data (e.g. [Francis & Francis, 1992](#); [Natanson *et al.*, 2001](#)).

LENGTH FREQUENCY ANALYSIS

Length frequency analysis subsumes a variety of different length-based methods, all of which produce estimates of growth rate. The corroboration occurs when the resulting growth estimate is compared to that of the age determination method. Monitoring of the progression of length frequency modes through time is one of the most basic of the length frequency analyses which is possible, and can be a reliable form of age corroboration in young, fast-growing fish (e.g. [Morales-Nin & Aldebert, 1997](#)). If monitoring occurs throughout the year, the results can be used to verify the annual frequency of the length modes, even if the corresponding age structures are not sampled. The subsequent comparisons of length at age or growth rate between length- and age-based methods must then be considered reasonably robust. Substantially less robust is the simple observation of the position of length modes in a single sample; in this approach, there is no confirmation that the modes actually correspond to any age classes, let alone the identity of that age class (e.g. [Shirvell,](#)

1981). At a more advanced level, methods such as Multifan fit a von Bertalanffy growth curve to multiple length frequency samples using maximum likelihood estimation (Fournier *et al.*, 1990). In instances where monthly length samples are available throughout the year, this method is a valuable approach for integrating multiple samples to produce estimates of growth rate (Francis & Francis, 1992). However, here as with the other length-based methods, the approach is most suited for young, fast-growing fish where the length modes for each age group are easily distinguished; Multifan will use the well-defined length modes of the younger fish to fit a growth model to all fish, even if the length frequency of the older fish is nonsensical. Size-selective migration into or out of the study area is not an allowable assumption of this, or any other length-based method.

SAMPLING OF RECRUITMENT PULSES

Periodic sampling of recruitment pulses has proven valuable in annulus studies of some long-lived fishes, although considerably less so at the daily level. Also termed 'progression of strong year-classes', the method compares the interval between periodic (e.g. yearly) samples and the increase in the apparent modal age of a recruitment pulse as determined through annulus counts. Where the recruitment pulses are sufficiently well-defined and there is no appreciable age-structured migration, mortality or age reader expectations (Beamish & McFarlane, 1995), the method can provide a strong, albeit qualitative, confirmation of growth increment periodicity (Donald *et al.*, 1992). For example, Morison *et al.* (1998a) clearly showed the otolith age-based modal progression of two strong year-classes over a 4 year sampling period, thus supporting the validity of the otolith-based ageing method. The method has also been used at the daily level (Uozumi & Ohara, 1993), although the likelihood of age-structured mortality distorting the apparent hatch date of the recruitment pulse is higher in young-of-the-year fishes (Campana & Jones, 1992). Comparison of the collection interval between samples and the difference in mean ages of those same samples is a variation on the same theme, but without the advantage of a well-defined recruitment pulse. Comparison of observed hatch dates with those estimated from the otolith microstructure of young-of-the-year (Morales-Nin *et al.*, 1999) is not a comparable measure of ageing accuracy, since it makes the difficult assumption of no age- or date-specific mortality.

NUMERICAL INTEGRATION OF DAILY INCREMENT WIDTHS

When the underlying assumptions are met, numerical integration of daily increment widths is an ingenious method for estimating (or even validating) annual age in species for which annular growth increment counts are problematic (Ralston & Miyamoto, 1983; Ralston & Williams, 1989; Smith & Kostlan, 1991). The method is based on a random sample of daily increment widths along an uninterrupted growth axis of the otolith which, when integrated over the observed length of the growth radius, must yield the daily age of the otolith and fish. However, the difficulty arises in assessing the validity of the underlying assumptions. If the daily increment sequence is anything other than continuous, or if the measured (and presumably clear) increment widths are not representative of the unmeasured (unclear) increments, the integration will fail, usually in

the form of underageing. While these assumptions have been ignored by some (Gauldie, 1994), it is difficult to overstate their importance. In general, daily increment sequences become increasingly likely to become interrupted in fish older than 1 year (Campana & Neilson, 1985), rendering the numerical integration method more useful for age corroboration than for validation.

DAILY INCREMENT COUNTS

Daily increment counts between presumed annuli can provide strong corroboration of the frequency of formation of the annuli, subject to the same assumptions of the numerical integration method. In this method however, all increments are examined and counted, making the assumptions of the approach somewhat easier to test. Daily increment formation must also be assumed, but this assumption is reasonably safe compared to that of sequence continuity. In cases where approximately 365 daily increments are counted between presumed annuli, a conclusion of annulus formation seems reasonably sound (e.g. Morales-Nin, 1988; Wilson, 1988). However, no conclusions can be drawn from otoliths in which markedly more or fewer increments are counted, since the result could reflect either the formation of non-annuli, misinterpretation of the daily increments, or an interrupted growth increment sequence.

ELEMENTAL AND ISOTOPIC CYCLES

Elemental and isotopic cycles have sometimes been observed in association with presumed annuli (Casselmann, 1983; Cailliet *et al.*, 1986; Stevenson & Secor, 1999), but as a form of age corroboration, they are purely correlative. Regular fluctuations in calcium, phosphorus or oxygen isotopes may well be reflecting environmental fluctuation, but whether they do so on an annual basis is subject to conjecture. It is generally accepted that annulus formation reflects annular variations in growth rate, but the frequent presence of pseudoannuli in many calcified structures indicates that those growth variations need not be annual. Since growth rate may also influence the deposition of some elements and isotopes (Campana, 1999), it is reasonable to expect that chemical cycles would mirror observed growth increments, whether or not those increments are formed annually. As a result, chemical cycles would appear useful for confirming the presence of visually-observed growth increments, but of limited value for inferring the periodicity of those increments.

OTHER METHODS

Methods which are neither validation nor corroboration are sometimes reported as such in the literature, but may actually serve some other purpose. For example, back-calculated lengths are useful for reducing the effect of size-selective sampling bias on the length estimates for the youngest fish in the sample, inferring the diameter of the first annulus, and providing a continuous sequence of lengths for a growth curve (e.g. Wintner & Cliff, 1999). However, the similarity of back-calculated lengths across several ageing structures in no way validates or corroborates any age interpretation; it merely shows consistency in the interpretation of the sequence of growth increments, independent of whether or not the interpretation is correct. Similarly, if the radius of a given growth increment is shown to form consistently at a particular 'age' (e.g.

Collins *et al.*, 1989; Morales-Nin, 2000), those results indicate that the increment is probably biologically meaningful (e.g. perhaps an annulus, perhaps a settling check), but its identity as an annulus remains unknown.

Comparison of multiple ageing structures within each fish is also a form of age non-corroboration. While structure comparisons are very useful during the selection of a preferred ageing method, consistency among within-fish growth structures is the rule rather than the exception. This is not surprising given that the growth of all structures within a given fish tends to be influenced by the same environmental and physiological factors.

VALIDATION OF THE FIRST GROWTH INCREMENT

Identification of the first, or innermost, growth increment is an important component of any age validation study. In studies which have validated increment periodicity rather than absolute age (as in chemical tagging studies), validation of the first increment is a mandatory adjunct to age determination; without a correctly defined starting point, age determinations will be consistently wrong by a constant amount. For example, uncertainty over the identification of the first vertebral annulus in porbeagle shark (*Lamna nasus* Bonnaterre) in two independent studies resulted in size at age estimates which consistently differed by one year (Francis & Stevens, 2000; Natanson *et al.*, 2001). Even in cases where absolute age has been validated, reliable identification of the first increment can substantially increase the precision of any individual age determination. Increased precision is particularly important for age determinations of young fish, where a random error of 1 year can introduce unacceptable error into all individual age estimates, even though the mean is still correct. Identification of the first annulus is often more problematic than that of the first daily growth increment, since the latter is often clearly visible as a hatch check (Campana & Neilson, 1985).

In principle, identification and validation of the first growth increment can proceed using any of the age validation methods described earlier. In practice though, only a subset of the available methods possess the necessary precision. Release of known age or chemically-marked young-of-the-year (YOY) fish is well suited to this type of application (Ferrell *et al.*, 1992; Fitzgerald *et al.*, 1997). For annulus studies, modal progression with age subsampling is also straight forward and accurate; monitoring the modal length of the presumed YOY (for example, through periodic research surveys) confirms their identity as YOY, while inspection of the marginal increment in those same YOY in the season of annulus formation confirms the formation of the presumed first annulus (Ferrell *et al.*, 1992). Since the marginal increment in YOY fish can sometimes be difficult to distinguish from false checks, a practical alternative is to measure the diameter of the YOY ageing structure (along the axis used for ageing) at the time of annulus formation to determine the expected diameter of the first annulus (Natanson *et al.*, 2001). Clearly, this approach requires independent knowledge of the season of annulus formation in that species and location, although that information can be gained from MIA of older fish (Beckman & Wilson, 1995). Validation of the first daily growth increment is even more straight forward,

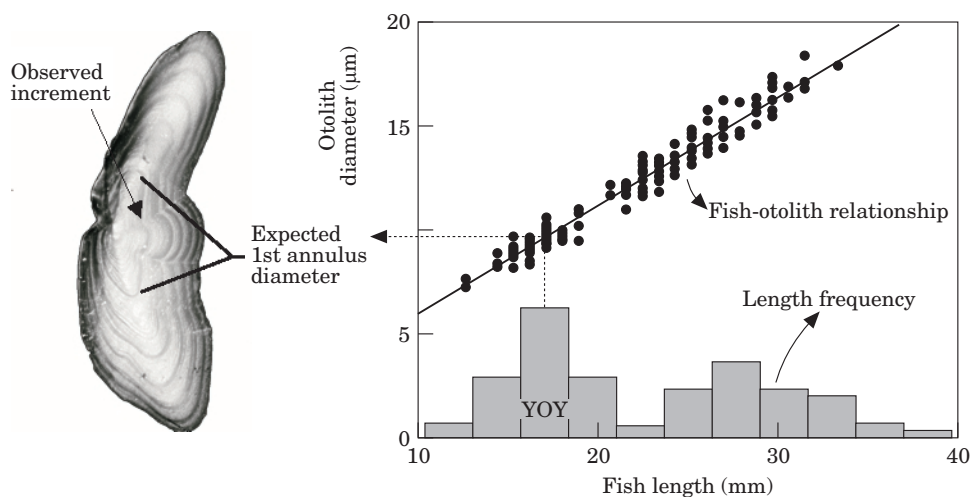


FIG. 3. Schematic diagram of an approach to validate the identity of the first annulus. Using the mode corresponding to the young-of-year (YOY) fish in a length-frequency sample collected around the time of annulus formation, annulus diameter is predicted on the basis of a fish length-otolith length regression (determined using the same or different samples). Predicted annulus diameter is then overlaid onto a series of otoliths to identify the growth increment which most closely corresponds in diameter.

requiring only samples of larvae hatched in captivity to identify the hatch or first-feeding check (Campana, 1989; Moksness, 1992).

In instances where ageing structures cannot be collected from YOY in the season when their first annulus is being formed, a variation on the modal progression validation approach is possible. This approach requires an estimate of mean YOY fish length in the season of annulus formation (e.g. around the first birthday). When this estimate is inserted into a fish length-otolith length regression, the mean expected diameter of the first annulus can be predicted (Fig. 3). It then becomes a simple matter to overlay the expected annulus diameter on probable first annuli in the otolith. This approach has been used to validate the position of the first annulus in both anchovy (*Sardinops sagax* Jenyns) (Spratt, 1975) and haddock otoliths (Campana, 1997).

In species with a clearly interpretable otolith microstructure, daily increment counts can often be used to confirm the identity of the first annulus (Waldron, 1994; Griffiths, 1996; Lehodey & Grandperrin, 1996). Since the increment counts need not be made with great accuracy, this is a robust approach as long as the approximate dates of hatch and annulus formation are known. The underlying assumption of an uninterrupted growth sequence must be satisfied, but daily increment formation is usually continuous during the first 6 months of life (Campana & Neilson, 1985). Since daily increments often become unresolvable during the winter months (Francis *et al.*, 1992), this method is most suited to the identification of the beginning of the winter growth zone, rather than its end.

Taking advantage of the fact that otoliths first appear in the embryonic stage of the fish, Morison *et al.* (1998a) used an otolith weight-age regression to determine if the first annulus had been correctly identified; a regression intercept

closest to zero was considered most consistent with accurate ageing. However, the basis for this method was not well established, and it was not at all clear that the method was sufficiently sensitive to detect anything other than gross errors.

INTERPRETATION OF THE MARGINAL GROWTH INCREMENT

Correct interpretation of the edge type influences ageing accuracy at the yearly level, since an annulus on the margin of a structure collected just after the assigned birthday can be given a different age assignment than the same structure collected just before the birthday. Although it has been suggested that edge type identification is a component of age validation (Francis *et al.*, 1992), the marginal increment does not require validation independent of that of any of the other growth increments. Thus it is more correct to say that the age assignment of a fish is a function not only of annulus count, but of edge type in relation to date of collection and assigned birth date. The protocols for handling edge type in relation to growth axis and collection date are more fully discussed elsewhere (Morison *et al.*, 1998a; Cappo *et al.*, 2000), as is the need to consider edge type on a stock-specific basis (Pearson, 1996). Casselman (1990) has automated the calculation of age from annulus count, edge type and collection date, so as to remove the possibility of calculation error from the age reader.

AGEING PRECISION

Precision is defined as the reproducibility of repeated measurements on a given structure, whether or not those measurements (age readings) are accurate. It is not unusual for inaccurate age readings to be highly reproducible (in other words, precisely wrong) or to show no relationship between accuracy and precision (Campana *et al.*, 1990; Campana & Moksness, 1991; Campana, 1995). 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 determining the age of a particular structure, of assessing the reproducibility of an individual's age determinations, or of comparing the skill level of one ager relative to that of others.

There are two widely used and statistically sound measures of ageing precision: average percent error (APE) and coefficient of variation (CV). Although percent agreement is the traditional index of ageing precision, many authors have pointed out its inadequacies (Beamish & Fournier, 1981; Chang, 1982; Kimura & Lyons, 1991; Campana *et al.*, 1995). The failure of percent agreement as a measure of precision is largely due to the fact that it varies so widely both among species and among ages within a species. Beamish & Fournier (1981) illustrated this point by noting that 95% agreement to within one year between two age readers of Pacific cod *Gadus macrocephalus* Tilesius constituted poor precision, given the few yearclasses in the fishery. On the other hand, 95% agreement to within 5 years would constitute good precision for spiny dogfish *Squalus acanthias* L., given its 60 year longevity. Thus, Beamish & Fournier (1981) recommended the use of average per cent error (APE), defined as:

$$\text{APE}_j = 100\% \times \frac{1}{R} \sum_{i=1}^R \frac{|x_{ij} - x_j|}{X_j}$$

where X_{ij} is the i th age determination of the j th fish, X_j is the mean age estimate of the j th fish, and R is the number of times each fish is aged. When averaged across many fish, it becomes an index of average percent error. Chang (1982) agreed that APE was a substantial improvement over percent agreement, but suggested that the standard deviation be substituted for the absolute deviation from the mean age. The resulting equation produces an estimate of the coefficient of variation (CV), expressed as the ratio of the standard deviation over the mean, and can be written as:

$$\text{CV}_j = 100\% \times \frac{\sqrt{\sum_{i=1}^R \frac{(X_{ij} - X_j)^2}{R-1}}}{X_j}$$

where CV_j is the age precision estimate for the j th fish. As with the equation for APE, it can be averaged across fish to produce a mean CV. The index of precision (D) is similar to the CV (and identical to APE when $R=2$), but is calculated as (Chang, 1982):

$$D_j = \frac{\text{CV}_j}{\sqrt{R}}$$

All measures of precision will be artificially inflated by any bias which exists between agers. In the absence of bias, the CV and APE are equally sensitive to precision differences among agers, although the CV is statistically more rigorous and thus more flexible (Chang, 1982). Neither APE nor CV is particularly sensitive to variations in age composition, although both tend to decline to asymptotic values as age increases. In contrast, percent agreement declines substantially with age, thus explaining the large artifactual variations in percent agreement which can occur when comparing the precision of two samples of different age composition (Kimura & Lyons, 1991; Campana *et al.*, 1995).

A review of 131 recent ageing papers reporting precision values indicates that both CV and APE are used widely, although most (57%) used CV. At the yearly level, the balance between CV and APE was almost exactly 50%, while CV was clearly favoured for otolith microstructure studies (84%). Such extensive use of two different indices has made it difficult to compare among many of the studies, despite the fact that Chang (1982) derived an equation indicating that CV will always exceed APE by a predictable quantity. To facilitate conversion between APE and CV, a predictive regression was fitted to the precision values reported in 14 papers in which both APE and CV were presented (Fig. 4). Three additional papers were not used, since one or both precision measures appear to have been calculated incorrectly. The regression demonstrated that CV is readily estimated from APE, that CV tends to be about 40% higher than APE for any given set of ageing data, and that the relationship between the two measures is

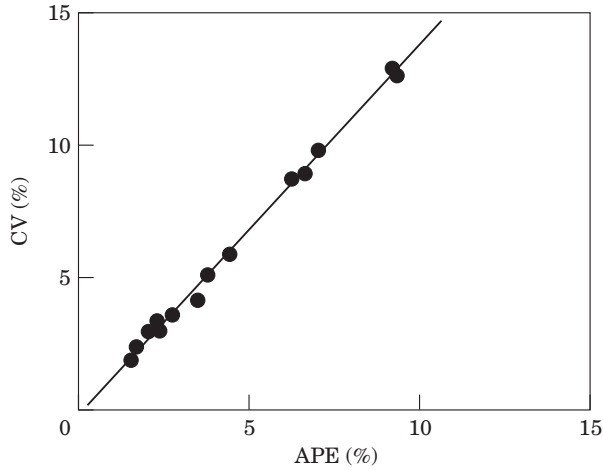


FIG. 4. Regression of coefficient of variation (CV) on average per cent error (APE) based on 14 papers which reported both measures of precision. $CV = -0.15 + 1.41 APE$.

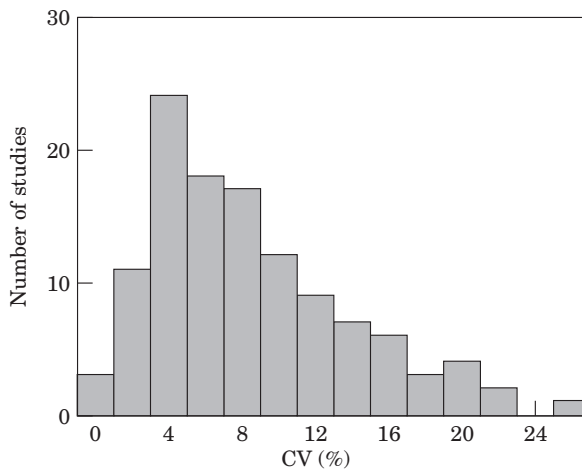


FIG. 5. Frequency histogram of published precision values ($n=117$) for both daily and annual age determinations across all ageing structures. Where precision was published in the form of APE, precision was converted to CV using the regression equation of Fig. 4.

tight. On the basis of this relationship, it is not self-evident that one measure is to be preferred over the other.

Published precision values ($n=117$) were summarized after first converting all to a common currency (CV) (Fig. 5). The median CV across all ageing structures, and including both annual and daily ageing studies, was 7.6%. The modal CV was 5%. To determine the precision levels characteristic of the most commonly aged structures, the CV values of Fig. 5 were broken down by structure within both daily and annual ageing studies (Fig. 6). There were no significant differences in precision among any of the ageing structures, although mean annual otolith ages tended to be slightly more precise than those of either scales or vertebrae. Mean precision was not well defined for fin rays or spines

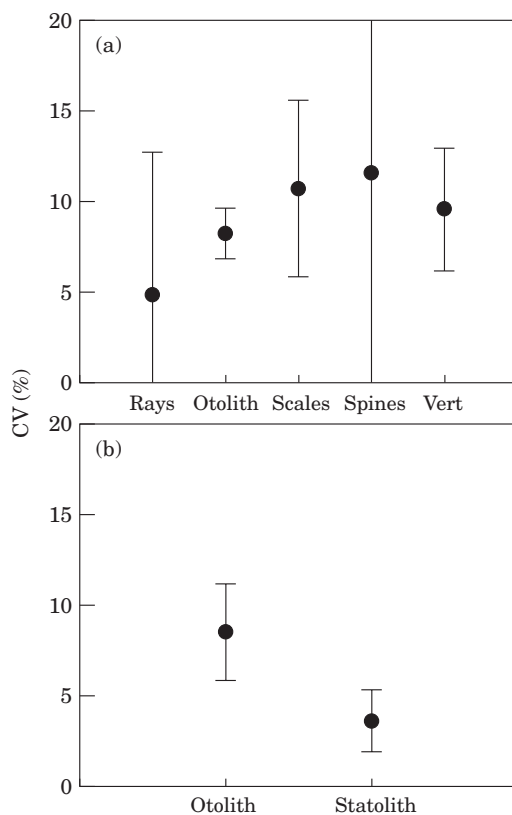


FIG. 6. Reported precision (calculated as CV) for the most frequently aged structures at both the annual (a) and daily (b) level. Error bars represent the mean \pm 95% confidence interval; vert, vertebrae.

due to low sample sizes. At the daily level, statoliths tended to be aged more precisely than otoliths. Interestingly, the CV of annual otolith ages was almost identical to that of daily otolith ages (Fig. 6).

There is no *a priori* value of precision which can be designated as a target level for ageing studies, since precision is highly influenced by the species and the nature of the structure, and not just the age reader. For example, virtually all studies reporting shark ages based on vertebrae did so with CV values exceeding 10%, while the most frequently reported CV for otoliths was 5%. High-volume ageing laboratories seldom report their threshold precision levels, but an APE less than 5% is expected in some (Morison *et al.*, 1998b). On the basis of the reviewed literature, many ageing studies can be carried out with a CV of less than 7.6%, corresponding to an APE of 5.5%. Informal discussions with a number of laboratories suggest that a CV of 5% serves as a reference point for many fishes of moderate longevity and reading complexity.

QUALITY CONTROL

In a review of ageing programs worldwide, Campana & Thorrold (2001) reported that 1–2 million fish were aged globally in 1999. The large majority of

these age determinations were made as part of 'production ageing' programs (*sensu* Morison *et al.*, 1998b), usually in support of fish stock assessments. Informal discussions with many of the production ageing laboratories suggested that the development and continued success of their ageing programs involved four steps:

- (1) development of an ageing method
- (2) age validation
- (3) preparation of a reference collection
- (4) quality control monitoring

The need for quality control monitoring (and reference collections) varies with the application, but is most relevant to high-volume and ongoing age determinations (Kimura & Lyons, 1991; Campana *et al.*, 1995; Morison *et al.*, 1998b). For these types of applications, ageing consistency from year to year is of prime importance, since even a gradual deterioration in ageing accuracy can lead to serious errors in a stock assessment (Lai & Gunderson, 1987; Tyler *et al.*, 1989; Bradford, 1991; Beamish & McFarlane, 1995; Campana, 1995; Eklund *et al.* 2000). Therefore, ageing consistency is monitored through time, under the working assumption that the method is accurate (an assumption that was presumably tested and confirmed in the age validation stage). As noted by Campana *et al.* (1995), 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 a set of statistical monitoring tools, both of which are discussed in the next sections.

REFERENCE COLLECTIONS

Reference collections of otoliths or other ageing structures are important, and perhaps mandatory, elements of an ongoing ageing program (Campana, 1995; Gröger, 1999). 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 testing for long-term drift, 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 insuring consistency in the type of structures which are interpreted as growth increments.

A reference collection can be defined as a collection of prepared ageing structures, of known or consensus-derived ages, representative of all factors which might reasonably be expected to influence the appearance or relative size of the growth increments. A list of such factors might include all combinations of age, sex, season, and source of collection, spanning the entire length range, a representative sample of the geographic range, and several collection years. Good seasonal coverage is necessary in light of the effect on the appearance of

the marginal increment. A collection composed of several years derived from across the distributional range is important to insure that any year-specific or region-specific anomalies are not given undue weight. In instances where stock differences can be expected, a stock-specific reference collection is warranted. A collection of optimal preparations is not required, and is in fact counterproductive: given its use for quality control monitoring and training, 'average' preparations are more likely to be of value than ideal preparations.

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 insures long-term availability, facilitates exchanges with other laboratories, and simplifies the training of new age readers. The use of annotated 'layers' (*sensu* Photoshop) which can be toggled off and on, allows the image to be interpreted with or without the annotation. Digital images can also be posted on the Web, thus helping to standardize age interpretations among other laboratories.

The exchange of either digital images or the actual reference collection presupposes that other laboratories use the same method of structure preparation. Mode of preparation can have a significant effect on age interpretation (Boehlert, 1985; Campana & Moksness, 1991), so it is inappropriate to consider that a reference collection prepared using sections (for example) will be optimal for a laboratory that uses only whole structures for ageing. There is no easy solution to this problem, other than for all participating laboratories to agree on a preferred method of preparation *a priori*. Although it is possible to prepare one otolith from each fish of the reference collection using one method, and the other otolith using an alternative method, this approach confounds preparation artifacts with interpretational differences, and cannot be recommended for the consensus ageing of a reference collection (although it is ideal as a test between preparation methods, and can certainly be used as an age bias test between laboratories). Since mode of preparation is not an issue for the development of a reference collection by a single laboratory, it need only be considered where multiple laboratories are sharing a single reference collection. There is, of course, no reason why multiple laboratories could not share a single reference collection if all used the same method of preparation. In the event that different methods are being used, the shared reference collection would have to be increased in size so as to include representative structures using all relevant methods.

To the extent that the ages assigned to the reference collection are accurate, all quality control monitoring using the collection will assess accuracy as well as consistency. Thus there is considerable value to insuring that the collection is aged as accurately as possible. Ideally, the ageing structures used in an age validation study will become part of the reference collection, since the ages of those structures are known with some confidence. In practice, known age structures are not always available. An excellent interim alternative is a collection which has been aged and annotated by international experts, perhaps as part of an otolith exchange program. An exchange program involving expert agers not only provides an initial comparison with the host age readers, it can provide the core of the reference collection at an early stage of the ageing program (Campana, 1995; Gröger, 1999). Clearly,

consensus-derived ages, even by experts, are not necessarily accurate. However, they are more likely to be so than those of an inexperienced ager, making the exchange of a reference collection an efficient and cost-effective approach of beginning an ageing program, with the basis for quality control already in place. The requirement for age validation is not removed however; in fact, validation becomes even more pressing once routine quality control monitoring is initiated.

The number of ageing structures required of a reference collection is somewhat arbitrary, but practical guidelines are possible. A minimum number for quality control monitoring would appear to be about 200, but even then, memorization of individual otoliths by experienced age readers is possible after several years. A number closer to 500 is preferable, where possible. If sent out as part of an exchange, it may be impractical to ask other laboratories to age the entire collection of 500, in which case a sub-sample of 200 or so could be exchanged. Structures used for training could make up a subset ($n \sim 100$) of the collection, carefully selected so as to demonstrate the key features. It is also important to note that the reference collection need not remain static through time; indeed, there is some advantage to adding a subsample of ageing structures (those used for monitoring tests each year) to the collection, thus allowing it to grow through time.

Permanent storage of otolith reference collections is important, but is not always possible. Otoliths are chemically stable if stored dry (Campana, 1999), but preparation for ageing may allow gradual deterioration through time. Mounted otolith thin sections appear to be durable, although some mounting media may crack or yellow with age. Charred otoliths may eventually lose contrast, although they can always be re-burnt (Chilton & Beamish, 1982). More problematic are whole otoliths stored in glycerine to enhance clearing. While over-clearing can sometimes be partially reversed, and while damage can often be delayed by removal of otoliths from the glycerine in between examinations, permanent storage of otoliths in glycerine often leads to irreversible clearing. For all modes of preparation, frequent handling of reference materials increases the probability of loss or damage. For these reasons alone, high quality digital images of each reference structure at an appropriate magnification is necessary to insure long-term availability of the collection. Many laboratories ageing large numbers of fish routinely use image analysis systems and digital images to age their samples (Planes *et al.*, 1991; Estep *et al.*, 1995; Macy, 1995; Morison *et al.*, 1998b). Therefore, imaging of ageing material is already an established procedure. In addition to insuring the permanence of the reference collection, the use of digital images simplifies training and facilitates exchange among laboratories.

QUALITY CONTROL MONITORING

Quality control (QC) monitoring can be defined as a process of inspection and measurement used to detect defects and deficiencies in a timely manner during the production of a product. In the case of age determinations, the product is the age interpretation, and the defects are ageing errors or inconsistencies in the manner in which the age is interpreted. Random errors are not usually a concern, since their extent is usually predictable and correctable using statistical

techniques (Richards *et al.*, 1992; Gröger, 1999). However, systematic or biased errors are of great concern, since they can lead to serious errors in population dynamics calculations (Lai & Gunderson, 1987; Rivard & Foy, 1987; Tyler *et al.*, 1989; Bradford, 1991). The gradual introduction of systematic ageing error over a period of years is in some ways even more serious, due to its more insidious nature (Campana, 1995). A quality control program monitors short- and long-term ageing consistency, both within and among age readers, by insuring that the age interpretation method does not drift through time. Thus, if the ageing method is accurate (validated), the monitoring will also assess accuracy.

Quality control programs associated with age determinations are in place in many laboratories, but have seldom been documented (Morison *et al.*, 1998b). Based on observation and discussion with many of these laboratories, the following QC protocol can be recommended:

- (1) early development of a reference collection, preferably consisting of known-age or consensus-aged structures;
- (2) periodic ageing of a randomly-drawn, blind-labelled subsample of the reference collection, intermixed with a subsample of structures recently aged as part of routine ageing. The combination of reference and recent samples insures that the age readers do not inadvertently change their ageing criteria during the QC test;
- (3) use of age bias graphs and CV as tools to evaluate the results of the monitoring (Campana *et al.*, 1995). Although simple to implement and interpret, this combination of graphics and statistics has proven effective in testing for both short- and long-term ageing consistency and precision.

In addition to its effectiveness in avoiding unexpected ageing problems, there are two cost-saving advantages to a QC program based on reference collections. The first advantage is that secondary age readers are not required. A secondary age reader may well be desired for contingency purposes, particularly if the continued availability of the primary age reader is in question. However, a secondary age reader is not needed to insure ageing accuracy or consistency, and may actually lead to an overly-optimistic view of ageing error (Heifetz *et al.*, 1999). Indeed, if discrepancies between primary and secondary age readers were to occur, comparisons with a reference collection would still be required to determine which of the age readers had drifted off course. A second advantage is that periodic exchanges of ageing structures with other laboratories are not required, aside from the first exchange used to develop ages for the reference collection. Once the consensus-derived ages for the reference collection are in place, there are no advantages to further exchanges, thus saving the time involved in implementing the exchanges.

When evaluating a QC monitoring test result for bias compared to the accepted reference ages, there are a variety of statistical tests and graphical measures available. Tests of symmetry (Hoenig *et al.*, 1995), ageing error matrices (Heifetz *et al.*, 1999), and a range of matched pair tests (Kimura & Lyons, 1991; Campana *et al.*, 1995) have all proven effective in detecting bias. The advantage of these tests is that they can be automated, warning the data manager of test results only if some threshold significance level has been reached.

Graphical measures cannot match the statistical rigour of the above tests, but can provide a more flexible and easily interpreted means of detecting and quantifying bias. *Campana et al. (1995)* noted that the age bias graph appears to be the best graphical measure of bias, since it provides an age by age measure of deviation away from an accepted, or reference, value. Therefore, it can clearly show under-ageing or over-ageing, even if the ageing error is restricted to the youngest or oldest fish. While a variety of statistical tests continue to be used, a review of the recent literature indicates that the age bias graph has become the most widely used measure of ageing bias.

Quality control monitoring based on a reference collection provides a level of error detection which cannot be matched by simple re-ageing of the previous years samples. To illustrate this point, the results of QC monitoring of Nova Scotian haddock just before the detection of serious ageing errors (*Campana, 1995*) is shown in *Fig. 7*. Based only on re-ageing of samples from the previous year, neither age bias graphs, statistical tests nor any measure of precision was capable of detecting the gradual, multi-year drift in age interpretation. Precision remained high, and there was no evidence of inconsistency between the age interpretations of the two years. However, when the same age reader aged the 'known age' reference collection, the ageing bias (which had developed over a period of seven years) became clearly evident (*Fig. 7*). Re-ageing of samples from the previous quarter or year is a common QC practice in many laboratories. However, this practice is essentially the same as measuring precision. The results shown here indicate that the practice of re-ageing a recent sample can be grossly inadequate for detecting a gradual shift in interpretation criteria (or accuracy), whether or not it is carried out by a second age reader. Use of a reference collection provides the stable reference point required for age comparisons, even in instances where multiple age readers gradually shift their ageing criteria in tandem. In instances where the reference collection is of known age, the test for consistency is also the test for accuracy.

There is little documentation available concerning the logistics of QC monitoring. Testing against a random subsample of a reference collection appears to be relatively common, as is the practice of including a subsample from the production ageing of the current or most recent year in order to insure that the reference collection is aged using the same criteria as was the production sample (*Campana, 1997; Heifetz et al., 1999*). If the reference and recent samples are randomly intermixed by a supervisor, preferably through use of digital images, it becomes very difficult for an age reader to consciously or unconsciously change his/her interpretation criteria for the QC test. 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. An additional advantage of this procedure is that the re-aged production sample will then have been aged twice, and can then be added to the reference collection for use in later years. Sample sizes of 100 from the reference collection plus 100 from a recent production run is sufficient to insure reasonable statistical power for the QC test.

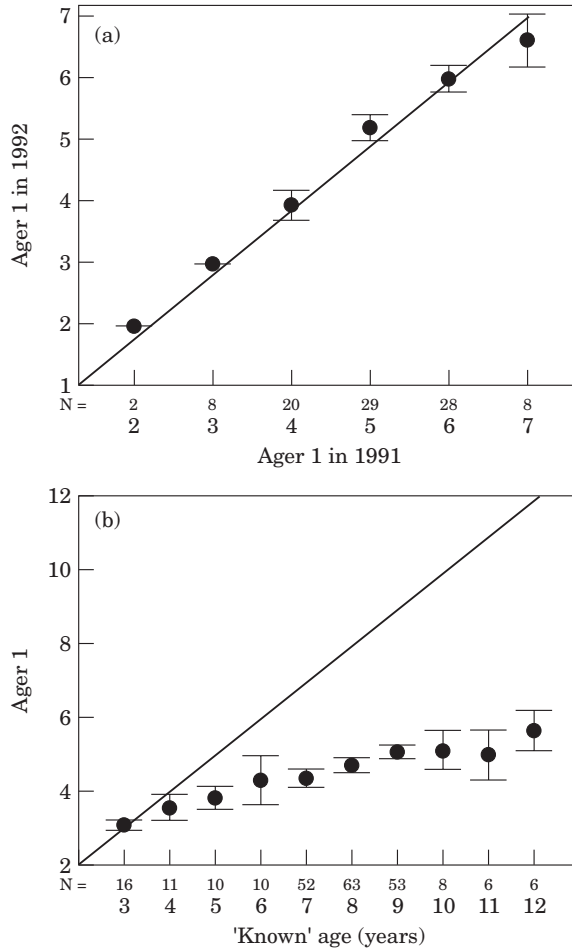


FIG. 7. Age bias graphs for haddock otoliths being aged in support of a stock assessment, just prior to the discovery that the ageing method was seriously underaging older fish. Extent of deviation of 95% confidence interval bars from 1 : 1 (—) line indicates extent of ageing bias. (a) Routine quality control test in which otoliths were re-aged 1 year later. No appreciable ageing bias is evident. Although this type of test is common in many laboratories, it was incapable of detecting the ageing problem, which developed over the course of 7 years; (b) Quality control test in which the same haddock ager read a reference collection aged by international experts. The large deviation from the 1:1 line indicates that haddock > age 5 were increasingly underaged by Ager 1, resulting in underaging by as much as 60% at age 12. (a) CV=3.4%; (b) CV=32.4%.

There is little advantage to carrying out more than one QC test per production run, since age readers are unlikely to change their ageing criteria in the middle of an ageing session. After a substantial interlude of non-ageing, or of ageing a different species, many experienced age readers re-orient themselves with the ageing of a species by re-reading a subsample of the previous year's production ageing. This is often done informally, and is not assessed as part of QC monitoring. QC monitoring is best done during or immediately after completion of production ageing. In instances where ageing is continuous, monitoring at a frequency of once per year is likely to be sufficient.

Identification of an 'ageing problem' from QC monitoring can be based on one or more threshold standards, but errors as large as those evident in Fig. 7 should never occur when proper quality control is in place. A change in precision will seldom be a reliable indicator of ageing error, particularly among experienced age readers (see Precision section). The age bias graph appears to be the most sensitive indicator of ageing bias, since it can be interpreted in terms of the type and magnitude of the bias, not just the presence/absence indicated by statistical tests. This is important, since not all bias warrants corrective action. For example, in the upper panel of Fig. 7, the 1992 set of age readings for ages 2–6 is higher than those for the same otoliths read in 1991; for some of the ages, the bias is statistically significant (e.g. the 95% confidence intervals do not cross the 1 : 1 line). Yet the magnitude of the bias, as indicated by the vertical interval between the 1 : 1 line and the data point, is relatively small ($< \frac{1}{2}$ year). The biological significance of such a small discrepancy is minimal. Similarly, if the mean for age 4 in the upper panel of Fig. 7 had happened to be 1 year above the 1 : 1 line, it is unlikely that any action would need to be taken; with no substantial bias evident for ages above or below age 4, the implications of a discrepancy at a single age are likely to be negligible. At what point does bias become important? A 1 year offset across a series of ages is certainly significant, since it indicates that one ager has counted an extra annulus, starting at the age at which the offset originated. A divergent trend between the data points and the 1 : 1 line, as shown in the lower panel of Fig. 7, is also troublesome, since it indicates that the two age readers are using different criteria to identify annuli. In the case of the lower panel of Fig. 7 for example, Ager 1 may have assumed that annulus width was constant and disregarded all narrower increments, whereas the international experts interpreted the progressively narrowing increments as annuli. In all cases however, an overlap of any individual error bar over the 1 : 1 line is an indication of the variance at that age, but is not a good criterion of the importance of any bias; for example, a 1 year offset seen at all ages would still indicate significant bias even if the error bars overlapped the 1 : 1 line at all ages.

If quality control monitoring identifies an ageing problem, the corrective action required will largely depend on the experience of the age reader and the number of structures which are being aged. In a production ageing environment, the following steps are often taken:

- (1) the ager re-reads the annotated training set to calibrate his/her interpretation against correctly identified growth increments;
- (2) conducts a blind test against a subsample of the reference collection, using an age bias graph to determine if the ageing bias is still present; if bias remains, returns to annotated training set for further training;
- (3) if bias is absent, re-ages all ageing structures from the point when the bias first appeared; in many cases, this point will not be known, requiring re-ageing from the time of the last QC test.

QUANTIFICATION AND CORRECTION OF AGEING ERROR

Ageing errors can be either random or biased, reflecting some combination of process and interpretation error. Although bias can be avoided through

validation studies and quality control, random error is virtually inevitable. Both forms of ageing error propagate through estimates of age at maturity, lifespan, population size and other vital rates, and due to their nonlinear relationships, can lead to biased results even when the underlying ageing error is random (Lai & Gunderson, 1987; Tyler *et al.*, 1989; Bradford, 1991). If the extent of the error can be quantified however, it can be corrected using statistical means.

Ageing error can be quantified using a variety of statistical models (Richards *et al.*, 1992; Heifetz *et al.*, 1999). Imprecision (random error) can be estimated on the basis of replicate readings of a given set of samples. However, all forms of ageing error can be quantified if the replicate age readings are from a known-age reference collection (Gröger, 1999; Heifetz *et al.*, 1999). In both cases, the product is an ageing error matrix, which can subsequently be used to statistically remove the ageing error from a set of age frequencies (Richards *et al.*, 1992).

Correction for ageing error is relatively straight forward when an unbiased ageing method has been used. The process is analogous to that of digital sharpening of an image. Figure 8 presents an example of a typical age frequency sample, aged with a commonly-observed level of precision (CV=7.6%). Statistical error correction removes the smoothing across age groups, thus amplifying most differences in age frequencies. The effect is most pronounced in older age groups, since ageing error at a given CV will spread an actual age across more age groups at an older age than at a younger age. It is for this reason that an age-structured population analysis will underestimate strong year-classes, and overestimate weak year-classes, in the absence of statistical correction of the catch at age data.

In principle, error correction of biased ageing data is possible (Gröger, 1999; Heifetz *et al.*, 1999). In practice, the reliability of the correction becomes increasingly questionable as the amount of bias increases. For example, the ageing bias evident in the lower panel of Fig. 7 indicates that the interpretations of the age reader have become virtually asymptotic after age 8. As a result, statistical correction of an independent set of aged samples would be almost impossible if the relative frequencies of each age group were different from that of the reference collection.

Age determinations, particularly the large-scale programs in support of stock assessments, are produced at a significant cost (Campana & Thorrold, 2001). While precision can always be improved through the ageing of larger sample sizes, the added costs must be balanced against the inevitable errors that will be introduced at other stages of the assessment process (e.g. accuracy of the catch values, randomness of the sampling, etc). Based on the analyses and reviews presented earlier, substantial cost savings are possible through implementation of proper quality control protocols, not the least of which is use of reference collections rather than secondary age readers. Use of statistical error correction for imprecision can reduce the number of ageing structures that need to be read. However, I would argue that the resulting cost savings should initially be reinvested into age validation in support of the reference collection. Such a one-time investment is very cost-effective given that it insures the long-term accuracy of the age data that is collected.

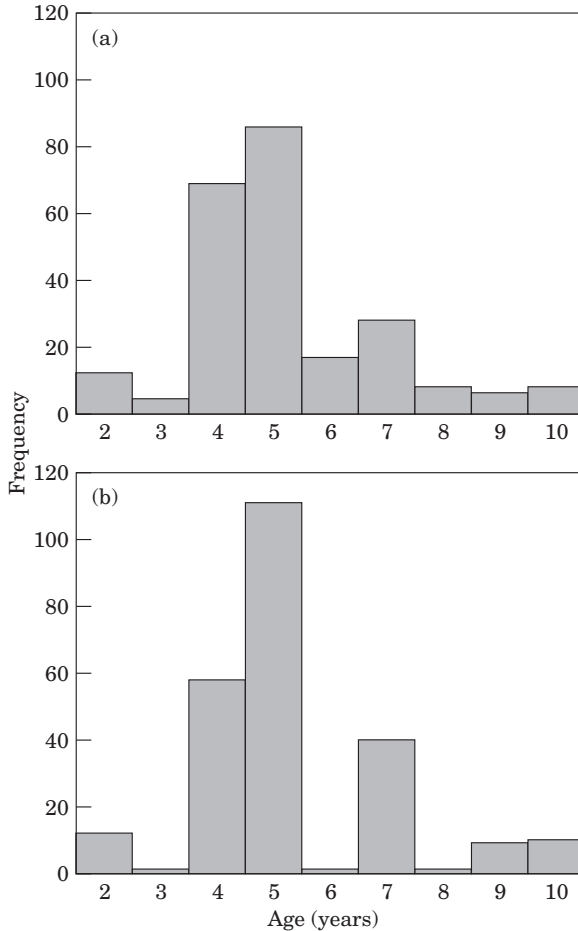


FIG. 8. Age composition of a hypothetical sample of fish before and after statistical removal of random ageing error. (a) Observed age composition after ageing with a CV of 7.6%, which is the median ageing error reported in all published ageing studies. (b) Actual age composition determined after statistical removal of random ageing error. Note the improved definition of the differences among age groups, removing the smoothing effect introduced by the ageing error. In particular, strongly represented age groups are made stronger, and poorly represented age groups are made weaker, especially in the older age groups.

I thank L. Marks and A. MacNeil for technical assistance, and two anonymous reviewers for helpful comments on the MS.

© 2001 Canadian Government

References

- Andrews, A. H., Coale, K. H., Nowicki, J. L., Lundstrom, C., Palacz, Z., Burton, E. J. & Cailliet, G. M. (1999). Application of an ion-exchange separation technique and thermal ionization mass spectrometry to ^{226}Ra determination in otoliths for radiometric age determination of long-lived fishes. *Canadian Journal of Fisheries and Aquatic Sciences* **56**, 1329–1338.

- Arkhipkin, A. I. (1997). Age and growth of the mesopelagic squid *Ancistrocheirus lesueurii* (Oegopsida: Ancistrocheiridae) from the central-east Atlantic based on statolith microstructure. *Marine Biology* **129**, 103–111.
- Bagenal, T. B. & Tesch, F. W. (1978). Age and growth. In *Methods for Assessment of Fish Production in Fresh Waters* (Bagenal, T., ed.), pp. 101–136. London: Blackwell Scientific.
- Baillon, N. (1992). Otolithometrie: revue et problèmes. In *Tissus durs et âge individuel des vertébrés* (Baglinière, J. L., Castanet, J., Conand, F. & Meunier, F. J., eds), pp. 21–52. Paris: ORSTOM Editions.
- Baker, A., Smart, P. L., Edwards, R. L. & Richards, D. A. (1993). Annual growth banding in a cave stalagmite. *Nature* **364**, 518–520.
- Baker, T. T. & Timmons, L. S. (1991). Precision of ages estimated from five bony structures of Arctic char (*Salvelinus alpinus*) from the Wood River System, Alaska. *Canadian Journal of Fisheries and Aquatic Sciences* **48**, 1007–1014.
- Beamish, R. J. (1979). New information on the longevity of Pacific ocean perch (*Sebastes alutus*). *Journal of the Fisheries Research Board of Canada* **36**, 1395–1400.
- Beamish, R. J. & Chilton, D. E. (1982). Preliminary evaluation of a method to determine the age of sablefish (*Anoplopoma fimbria*). *Canadian Journal of Fisheries and Aquatic Sciences* **39**, 277–287.
- Beamish, R. J. & Fournier, D. A. (1981). A method for comparing the precision of a set of age determinations. *Canadian Journal of Fisheries and Aquatic Sciences* **38**, 982–983.
- Beamish, R. J. & McFarlane, G. A. (1983). The forgotten requirement for age validation in fisheries biology. *Transactions of the American Fisheries Society* **112**, 735–743.
- Beamish, R. J. & McFarlane, G. A. (1985). Annulus development on the second dorsal spine of the spiny dogfish (*Squalus acanthias*) and its validity for age determination. *Canadian Journal of Fisheries and Aquatic Sciences* **42**, 1799–1805.
- Beamish, R. J. & McFarlane, G. A. (1987). Current trends in age determination methodology. In *Age and Growth of Fish* (Summerfelt, R. C. & Hall, G. E., eds), pp. 15–42. Ames: Iowa State University Press.
- Beamish, R. J. & McFarlane, G. A. (1995). A discussion of the importance of aging errors, and an application to walleye pollock: the world's largest fishery. In *Recent Developments in Fish Otolith Research* (Secor, D. H., Dean, J. M. & Campana, S. E., eds), pp. 545–565. Columbia: University of South Carolina Press.
- Beamish, R. J. & McFarlane, G. A. (2000). Reevaluation of the interpretation of annuli from otoliths of a long-lived fish, *Anoplopoma fimbria*. *Fisheries Research* **46**, 105–111.
- Beauchemin, D., McLaren, J. W., Mykytiuk, A. P. & Berman, S. S. (1987). Determination of trace metals in a river water reference material by inductively coupled plasma mass spectrometry. *Analytical Chemistry* **59**, 778–783.
- Beckman, D. W. & Wilson, C. A. (1995). Seasonal timing of opaque zone formation in fish otoliths. In *Recent Developments in Fish Otolith Research* (Secor, D. H., Dean, J. M. & Campana, S. E., eds), pp. 27–43. Columbia: University of South Carolina Press.
- Bennett, J. T., Boehlert, G. W. & Turekian, K. K. (1982). Confirmation of longevity in *Sebastes diploproa* (Pisces: Scorpaenidae) from Pb-210/Ra-226 measurements in otoliths. *Marine Biology* **71**, 209–215.
- Blacker, R. W. (1974). Recent advances in otolith studies. In *Sea Fisheries Research* (Harden Jones, F. R., ed.), pp. 67–90. London: Elek Science.
- Boehlert, G. W. (1985). Using objective criteria and multiple regression models for age determination in fishes. *Fishery Bulletin, U.S.* **83**, 103–117.
- Bradford, M. J. (1991). Effects of ageing errors on recruitment time series estimated from sequential population analysis. *Canadian Journal of Fisheries and Aquatic Sciences* **48**, 555–558.
- Brothers, E. B. (1979). Age and growth studies on tropical fishes. In *Stock Assessment for Tropical Small-Scale Fisheries* (Saila, S. B. & Roedel, P. M., eds), pp. 119–136. Kingston: University of Rhode Island.

- Brown, C. A. & Gruber, S. H. (1988). Age assessment of the lemon shark, *Negaprion brevirostris*, using tetracycline validated vertebral centra. *Copeia* **1988**, 747–753.
- Burton, E. J., Andrews, A. H., Coale, K. H. & Cailliet, G. M. (1999). Application of radiometric age determination to three long-lived fishes using ^{210}Pb : ^{226}Ra disequilibria in calcified structures: a review. *American Fisheries Society Symposium* **23**, 77–87.
- Cailliet, G. M. (1990). Elasmobranch age determination and verification: an updated review. *National Oceanic and Atmospheric Administration Technical Report* **90**, 157–165.
- Cailliet, G. M., Radtke, R. L. & Weldon, B. A. (1986). Elasmobranch age determination and verification: a review. In *Indo-Pacific Fish Biology: Proceedings of the Second International Conference on Indo-Pacific Fishes* (Uyeno, T., Arai, R., Taniuchi, T. & Matsuura, K., eds), pp. 345–360. Tokyo: Ichthyological Society of Japan.
- Campana, S. E. (1983). Feeding periodicity and the production of daily growth increments in otoliths of steelhead trout (*Salmo gairdneri*) and starry flounder (*Platichthys stellatus*). *Canadian Journal of Zoology* **61**, 1591–1597.
- Campana, S. E. (1984). Comparison of age determination methods for the starry flounder. *Transactions of the American Fisheries Society* **113**, 365–369.
- Campana, S. E. (1989). Otolith microstructure of three larval gadids in the Gulf of Maine, with inferences on early life history. *Canadian Journal of Zoology* **67**, 1401–1410.
- Campana, S. (1995). Expert age determination of 4VW and 4X haddock otoliths by national and international laboratories. *Department of Fisheries and Oceans Atlantic Fisheries Research Document* **95/120**.
- Campana, S. E. (1997). Use of radiocarbon from nuclear fallout as a dated marker in the otoliths of haddock, *Melanogrammus aeglefinus*. *Marine Ecology Progress Series* **150**, 49–56.
- Campana, S. E. (1999). Chemistry and composition of fish otoliths: pathways, mechanisms and applications. *Marine Ecology Progress Series* **188**, 263–297.
- Campana, S. E. & Jones, C. M. (1992). Analysis of otolith microstructure data. In *Otolith microstructure examination and analysis* (Stevenson, D. K. & Campana, S. E., eds), pp. 73–100. Canadian Special Publication of Fisheries and Aquatic Sciences **117**.
- Campana, S. E. & Jones, C. M. (1998). Radiocarbon from nuclear testing applied to age validation of black drum, *Pogonias cromis*. *Fishery Bulletin, U.S.* **96**, 185–192.
- Campana, S. E. & Moksness, E. (1991). Accuracy and precision of age and hatch date estimates from otolith microstructure examination. *ICES Journal of Marine Science* **48**, 303–316.
- Campana, S. E. & Neilson, J. D. (1982). Daily growth increments in otoliths of starry flounder (*Platichthys stellatus*) and the influence of some environmental variables in their production. *Canadian Journal of Fisheries and Aquatic Sciences* **39**, 937–942.
- Campana, S. E. & Neilson, J. D. (1985). Microstructure of fish otoliths. *Canadian Journal of Fisheries and Aquatic Sciences* **42**, 1014–1032.
- Campana, S. E. & Thorrold, S. R. (2001). Otoliths, increments and elements: keys to a comprehensive understanding of fish populations? *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 30–38.
- Campana, S. E., Zwanenburg, K. C. T. & Smith, J. N. (1990). $^{210}\text{Pb}/^{226}\text{Ra}$ determination of longevity in redfish. *Canadian Journal of Fisheries and Aquatic Sciences* **47**, 163–165.
- Campana, S. E., Oxenford, H. A. & Smith, J. N. (1993). Radiochemical determination of longevity in flyingfish *Hirundichthys affinis* using Th-228/Ra-228. *Marine Ecology Progress Series* **100**, 211–219.
- Campana, S. E., Annand, M. C. & McMillan, J. I. (1995). Graphical and statistical methods for determining the consistency of age determinations. *Transactions of the American Fisheries Society* **124**, 131–138.

- Cappo, M., Eden, P., Newman, S. J. & Robertson, S. (2000). A new approach to validation of periodicity and timing of opaque zone formation in the otoliths of eleven species of *Lutjanus* from the central Great Barrier Reef. *Fishery Bulletin, U.S.* **98**, 474–488.
- Carlson, J. K., Cortés, E. & Johnson, A. G. (1999). Age and growth of the blacknose shark, *Carcharhinus acronotus*, in the eastern Gulf of Mexico. *Copeia* **1999**, 684–691.
- Cass, A. J. & Beamish, R. J. (1983). First evidence of validity of the fin-ray method of age determination for marine fishes. *North American Journal of Fisheries Management* **3**, 182–188.
- Casselman, J. M. (1974). Analysis of hard tissue of pike *Esox lucius* with special reference to age and growth. In *The Ageing of Fish* (Bagenal, T. B., ed.), pp. 13–27. Surrey: Unwin Brothers.
- Casselman, J. M. (1983). Age and growth assesment of fish from their calcified structures—techniques and tools. *National Oceanic and Atmospheric Administration Technical Report NMFS* **8**, 1–17.
- Casselman, J. M. (1987). Determination of age and growth. In *The Biology of Fish Growth* (Weatherley, A. H. & Gill, H. S., eds), pp. 209–242. New York: Academic Press.
- Casselman, J. M. (1990). Growth and relative size of calcified structures of fish. *Transactions of the American Fisheries Society* **119**, 673–688.
- Chang, W. Y. B. (1982). A statistical method for evaluating the reproducibility of age determination. *Canadian Journal of Fisheries and Aquatic Sciences* **39**, 1208–1210.
- Chilton, D. E. & Beamish, R. J. (1982). Age determination methods for fishes studied by the Groundfish Program at the Pacific Biological Station. *Canadian Special Publication of Fisheries and Aquatic Sciences* **60**, 1–102.
- Chivas, A. R., De Deckker, P. & Shelley, J. M. G. (1985). Strontium content of ostracods indicates lacustrine palaeosalinity. *Nature* **316**, 251–253.
- Collins, M. R., Schmidt, D. J., Waltz, C. W. & Pickney, J. L. (1989). Age and growth of king mackerel, *Scomberomorus cavalla*, from the Atlantic coast of the United States. *Fishery Bulletin, U.S.* **87**, 49–61.
- Dodge, R. E. & Thomson, J. (1974). The natural radiochemical and growth records in contemporary hermatypic corals from the Atlantic and Caribbean. *Earth and Planetary Science Letters* **23**, 313–322.
- Donald, D. B., Babaluk, J. A., Craig, J. F. & Musker, W. A. (1992). Evaluation of the scale and operculum methods to determine age of adult goldeyes with special reference to a dominant year-class. *Transactions of the American Fisheries Society* **121**, 792–796.
- Eklund, J., Parmanne, R. & Aneer, G. (2000). Between-reader variation in herring otolith ages and effects on estimated population parameters. *Fisheries Research* **46**, 147–154.
- Estep, K. W., Nedreaas, K. H. & MacIntyre, F. (1995). Computer image enhancement and presentation of otoliths. In *Recent Developments in Fish Otolith Research* (Secor, D. H., Dean, J. M. & Campana, S. E., eds), pp. 303–317. Columbia: University of South Carolina Press.
- Fenton, G. E., Ritz, D. A. & Short, S. A. (1990). 210-Pb/226-Ra disequilibria in otoliths of blue grenadier, *Macruronus novaezelandiae*: problems associated with radiometric ageing. *Australian Journal of Marine and Freshwater Research* **41**, 467–473.
- Fenton, G. E., Ritz, D. A. & Short, S. A. (1991). Age determination of orange roughy, *Hoplostethus atlanticus* (Pisces: Trachichthyidae) using Pb-210:Ra-226 disequilibria. *Marine Biology* **109**, 197–202.
- Ferrell, D. J., Henry, G. W., Bell, J. D. & Quartararo, N. (1992). Validation of annual marks in the otoliths of young snapper, *Pagrus auratus* (Sparidae). *Australian Journal of Marine and Freshwater Research* **43**, 1051–1055.

- Fitzgerald, T. J., Margenau, T. L. & Copes, F. A. (1997). Muskellunge scale interpretation: the question of aging accuracy. *North American Journal of Fisheries Management* **17**, 206–209.
- Folkvord, A., Rukan, K., Johannessen, A. & Moksness, E. (1997). Early life history of herring larvae in contrasting feeding environments determined by otolith microstructure analysis. *Journal of Fish Biology* **51**(Suppl. A), 250–263.
- Foreman, T. (1987). A method of simultaneously tagging large oceanic fish and injecting them with tetracycline. *Fishery Bulletin, U.S.* **85**, 645–647.
- Fournier, D. A., Sibert, J. R., Majkowski, J. & Hampton, J. (1990). MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Canadian Journal of Fisheries and Aquatic Sciences* **47**, 301–317.
- Fowler, A. J. (1990). Validation of annual growth increments in the otoliths of a small, tropical coral reef fish. *Marine Ecology Progress Series* **64**, 25–38.
- Fowler, A. J. & Short, D. A. (1998). Validation of age determination from otoliths of the King George whiting *Sillaginoides punctata* (Perciformes). *Marine Biology* **130**, 577–587.
- Francis, M. P. & Francis, R. I. C. C. (1992). Growth rate estimates for New Zealand rig (*Mustelus lenticulatus*). *Australian Journal of Marine and Freshwater Research* **43**, 1157–1176.
- Francis, M. P. & Stevens, J. D. (2000). Reproduction, embryonic development and growth of the porbeagle shark, *Lamna nasus*, in the southwest Pacific Ocean. *Fishery Bulletin, U.S.* **98**, 41–63.
- Francis, R. I. C. C. (1988). Are growth parameters estimated from tagging and age-length data comparable? *Canadian Journal of Fisheries and Aquatic Sciences* **45**, 936–942.
- Francis, R. I. C. C. & Horn, P. L. (1997). Transition zone in otoliths of orange roughy (*Hoplostethus atlanticus*) and its relationship to the onset of maturity. *Marine Biology* **129**, 681–687.
- Francis, R. I. C. C., Paul, L. J. & Mulligan, K. P. (1992). Ageing of adult snapper (*Pagrus auratus*) from otolith annual ring counts: validation by tagging and oxytetracycline injection. *Australian Journal of Marine and Freshwater Research* **43**, 1069–1089.
- Gage, J. D. (1990). Skeletal growth markers in the deep-sea brittle stars *Ophiura ljungmani* and *Ophiomusium lymani*. *Marine Biology* **104**, 427–435.
- Gauldie, R. W. (1994). The morphological basis of fish age estimation methods based on the otolith of *Nemadactylus macropterus*. *Canadian Journal of Fisheries and Aquatic Sciences* **51**, 2341–2362.
- Gauldie, R. W., Coote, G. & West, I. F. (1993). The otoliths of senescent kahawai, *Arripis trutta* (Arripidae). *Cybius* **17**, 25–37.
- Geffen, A. J. (1982). Otolith ring deposition in relation to growth rate in herring (*Clupea harengus*) and turbot (*Scophthalmus maximus*) larvae. *Marine Biology* **71**, 317–326.
- Geffen, A. J. (1987). Methods of validating daily increment deposition in otoliths of larval fish. In *Age and Growth of Fish* (Summerfelt, R. C. & Hall, G. E., eds), pp. 123–240. Ames: Iowa State University Press.
- Geffen, A. J. (1992). Validation of otolith increment deposition rate. In *Otolith Microstructure Examination and Analysis* (Stevenson, D. K. & Campana, S. E., eds), pp. 101–113. Canadian Special Publication of Fisheries and Aquatic Sciences **117**.
- Germano, D. J. (1998). Scutes and age determination of desert tortoises revisited. *Copeia* **1998**, 482–484.
- Goodfriend, G. A. (1992). Rapid racemization of aspartic acid in mollusc shells and potential for dating over recent centuries. *Nature* **357**, 399–401.
- Goren, A. D., Brodie, P. F., Spotte, S., Ray, G. C., Kaufman, H. W., Gwinnett, A. J., Sciubba, J. J. & Buck, J. D. (1987). Growth layer groups (GLGs) in the teeth of

- an adult belukha whale (*Delphinapterus leucas*) of known age: evidence for two annual layers. *Marine Mammal Science* **3**, 14–21.
- Griffiths, M. H. (1996). Age and growth of South African silver kob *Argyrosomus inodorus* (Sciaenidae), with evidence for separate stocks. *South African Journal of Marine Science* **17**, 37–48.
- Gröger, J. (1999). A theoretical note on the interpersonal correction of age readings by means of calibration techniques. *Archive of Fishery and Marine Research* **47**, 77–101.
- Gulland, J. A. & Holt, S. J. (1959). Estimation of growth parameters for data at unequal time intervals. *Journal de Conseil Conseil International pour l'Exploration de la Mer* **25**, 47–49.
- Hammer, C. & Braum, E. (1988). Quantification of age pigments (lipofuscin). *Comparative Biochemistry and Physiology* **90B**, 7–17.
- Hanchet, S. M. & Uozumi, Y. (1996). Age validation and growth of southern blue whiting, *Micromesistius australis*, in New Zealand. *New Zealand Journal of Marine and Freshwater Research* **30**, 57–67.
- Heifetz, J., Anderl, D., Maloney, N. E. & Rutecki, T. L. (1999). Age validation and analysis of ageing error from marked and recaptured sablefish, *Anoplopoma fimbria*. *Fishery Bulletin, U.S.* **97**, 256–263.
- Hoening, J. M., Morgan, M. J. & Brown, C. A. (1995). Analysing differences between two age determination methods by tests of symmetry. *Canadian Journal of Fisheries and Aquatic Sciences* **52**, 364–368.
- Holmden, C., Creaser, R. A. & Muehlenbachs, K. (1997). Paleosalinities in ancient brackish water systems determined by Sr-87/Sr-86 ratios in carbonate fossils: a case study from the western Canada sedimentary basin. *Geochimica Cosmochimica Acta* **61**, 2105–2118.
- Hyndes, G. A., Loneragan, N. R. & Potter, I. C. (1992). Influence of sectioning otoliths on marginal increment trends and age and growth estimates for the flathead *Platycephalus speculator*. *Fishery Bulletin, U.S.* **90**, 276–284.
- Jenkins, G. P. & Davis, T. L. O. (1990). Age, growth rate, and growth trajectory determined from otolith microstructure of southern bluefin tuna *Thunnus maccoyii* larvae. *Marine Ecology Progress Series* **63**, 93–104.
- Jones, C. (1986). Determining age of larval fish with the otolith increment technique. *Fishery Bulletin, U.S.* **84**, 91–104.
- Kalish, J. M. (1993). Pre- and post-bomb radiocarbon in fish otoliths. *Earth and Planetary Science Letters* **114**, 549–554.
- Kalish, J. M. (1995a). Application of the bomb radiocarbon chronometer to the validation of redfish *Centroberyx affinis* age. *Canadian Journal of Fisheries and Aquatic Sciences* **52**, 1399–1405.
- Kalish, J. M. (1995b). Radiocarbon and fish biology. In *Recent Developments in Fish Otolith Research* (Secor, D. H., Dean, J. M. & Campana, S. E., eds), pp. 637–653. Columbia: University of South Carolina Press.
- Kalish, J. M., Beamish, R. J., Brothers, E. B., Casselman, J. M., Francis, R. I. C. C., Mosegaard, H., Panfili, J., Prince, E. D., Thresher, R. E., Wilson, C. A. & Wright, P. J. (1995). Glossary for otolith studies. In *Recent Developments in Fish Otolith Research* (Secor, D. H., Dean, J. M. & Campana, S. E., eds), pp. 723–729. Columbia: University of South Carolina Press.
- Kalish, J. M., Johnston, J. M., Gunn, J. S. & Clear, N. P. (1996). Use of the bomb radiocarbon chronometer to determine age of southern bluefin tuna *Thunnus maccoyii*. *Marine Ecology Progress Series* **143**, 1–8.
- Kalish, J. M., Johnston, J. M., Smith, D. C., Morison, A. K. & Robertson, S. G. (1997). Use of the bomb radiocarbon chronometer for age validation in the blue grenadier *Macruronus novaezelandiae*. *Marine Biology* **128**, 557–563.
- Kastelle, C. R., Kimura, D. K., Nevissi, A. E. & Gunderson, D. R. (1994). Using Pb-210/Ra-226 disequilibria for sablefish, *Anoplopoma fimbria*, age validation. *Fishery Bulletin, U.S.* **92**, 292–301.

- Kimura, D. K. & Lyons, J. J. (1991). Between-reader bias and variability in the age-determination process. *Fishery Bulletin, U.S.* **89**, 53–60.
- Kuniholm, P. I., Kromer, B., Manning, S. W., Newton, M., Latini, C. E. & Bruce, M. J. (1996). Anatolian tree rings and the absolute chronology of the eastern Mediterranean, 2220–718 BC. *Nature* **381**, 780–783.
- Labropoulou, M. & Papaconstantinou, C. (2000). Comparison of otolith growth and somatic growth in two macrourid fishes. *Fisheries Research* **46**, 177–188.
- Lai, H. L. & Gunderson, D. R. (1987). Effects of ageing errors on estimates of growth, mortality and yield per recruit for walleye pollock (*Theragra chalcogramma*). *Fisheries Research* **5**, 287–302.
- Laur, R. M., Nishimoto, R. & Wetherall, J. A. (1985). Frequency of increment formation on sagittae of north Pacific albacore (*Thunnus alalunga*). *Canadian Journal of Fisheries and Aquatic Sciences* **42**, 1552–1555.
- Lee, D. W. & Prince, E. D. (1995). Analysis of otoliths and vertebrae from nine tag-recaptured Atlantic bluefin tuna (*Thunnus thynnus*). In *Recent Developments in Fish Otolith Research* (Secor, D. H., Dean, J. M. & Campana, S. E., eds), pp. 361–374. Columbia: University of South Carolina Press.
- Lehodey, P. & Grandperrin, R. (1996). Age and growth of the alfonsino *Beryx splendens* over the seamounts off New Caledonia. *Marine Biology* **125**, 249–258.
- Lowerre-Barbieri, S. K., Chittenden, M. E. Jr & Jones, C. M. (1994). A comparison of a validated otolith method to age weakfish, *Cynoscion regalis*, with the traditional scale method. *Fishery Bulletin, U.S.* **92**, 555–568.
- Lutz, R. A. & Rhoads, D. C. (1980). Growth patterns within the molluscan shell: an overview. In *Skeletal Growth of Aquatic Organisms* (Rhoads, D. C. & Lutz, R. A., eds), pp. 203–254. New York: Plenum Press.
- MacLellan, S. E. & Saunders, M. W. (1995). A natural tag on the otoliths of Pacific hake (*Merluccius productus*) with implications for age validation and migration. In *Recent Developments in Fish Otolith Research* (Secor, D. H., Dean, J. M. & Campana, S. E., eds), pp. 567–580. Columbia: University of South Carolina Press.
- McFarlane, G. A. & Beamish, R. J. (1995). Validation of the otolith cross-section method of age determination for sablefish (*Anoplopoma fimbria*) using oxytetracycline. In *Recent Developments in Fish Otolith Research* (Secor, D. H., Dean, J. M. & Campana, S. E., eds), pp. 319–329. Columbia: University of South Carolina Press.
- Macy III, W. K. (1995). The application of digital image processing to the aging of long-finned squid, *Loligo pealei*, using the statolith. In *Recent Developments in Fish Otolith Research* (Secor, D. H., Dean, J. M. & Campana, S. E., eds), pp. 283–302. Columbia: University of South Carolina Press.
- Marshall, S. L. & Parker, S. S. (1982). Pattern identification in the microstructure of sockeye salmon (*Oncorhynchus nerka*) otoliths. *Canadian Journal of Fisheries and Aquatic Sciences* **39**, 542–547.
- Matlock, G. C., Colura, R. L. & McEachron, L. W. (1993). Direct validation of black drum (*Pogonias cromis*) ages determined from scales. *Fishery Bulletin, U.S.* **91**, 558–563.
- Meekan, M. G. & Fortier, L. (1996). Selection for fast growth during the larval life of Atlantic cod *Gadus morhua* on the Scotian Shelf. *Marine Ecology Progress Series* **137**, 25–37.
- Milton, D. A., Short, S. A., O'Neill, M. F. & Blaber, S. J. M. (1995). Ageing of three species of tropical snapper (Lutjanidae) from the Gulf of Carpentaria, Australia, using radiometry and otolith ring counts. *Fishery Bulletin, U.S.* **93**, 103–115.
- Moksness, E. (1992). Validation of daily increments in the otolith microstructure of Norwegian spring-spawning herring (*Clupea harengus*). *ICES Journal of Marine Science* **49**, 231–235.
- Morales-Nin, B. (1988). Age determination in a tropical fish, *Lethrinus nebulosus* (Forsk., 1775) (Teleostei: Lethrinidae) by means of otolith interpretation. *Investigaciones Pesquera* **52**, 237–244.

- Morales-Nin, B. (1989). Growth determination of tropical marine fishes by means of otolith interpretation and length frequency analysis. *Aquatic Living Resources* **2**, 241–253.
- Morales-Nin, B. (2000). Review of the growth regulation processes of otolith daily increment formation. *Fisheries Research* **46**, 53–67.
- Morales-Nin, B. & Aldebert, Y. (1997). Growth of juvenile *Merluccius merluccius* in the Gulf of Lions (NW Mediterranean) based on otolith microstructure and length-frequency analysis. *Fisheries Research* **30**, 77–85.
- Morales-Nin, B., Torres, G. J., Lombarte, A. & Recasens, L. (1998). Otolith growth and age estimation in the European hake. *Journal of Fish Biology* **53**, 1155–1168.
- Morales-Nin, B., di Stefano, M., Potoschi, A., Massuti, E., Rizzo, P. & Gancitano, S. (1999). Differences between the sagitta, lapillus and vertebra in estimating age and growth in juvenile Mediterranean dolphinfish (*Coryphaena hippurus*). *Scientia Marina* **63**, 327–336.
- Morison, A. K., Coutin, P. C. & Robertson, S. G. (1998a). Age determination of black bream, *Acanthopagrus butcheri* (Sparidae), from the Gippsland Lakes of south-eastern Australia indicates slow growth and episodic recruitment. *Marine and Freshwater Research* **49**, 491–498.
- Morison, A. K., Robertson, S. G. & Smith, D. G. (1998b). An integrated system for production fish aging: image analysis and quality assurance. *North American Journal of Fisheries Management* **18**, 587–598.
- Natanson, L. J., Casey, J. G., Kohler, N. E. & Colket IV, T. (1999). Growth of the tiger shark, *Galeocerdo cuvier*, in the western North Atlantic based on tag returns and length frequencies, and a note on the effects of tagging. *Fishery Bulletin, U.S.* **97**, 944–953.
- Natanson, L. J., Mello, J. J. & Campana, S. E. (2001). Validated age and growth of the porbeagle shark, *Lamna nasus*, in the western North Atlantic Ocean. *Fishery Bulletin, U.S.*, in press.
- Nedreaas, K. (1990). Age determination of Northeast Atlantic *Sebastes* species. *Journal de Conseil Conseil International pour l'Exploration de la Mer* **47**, 208–230.
- Olive, P. J. W. (1980). Growth lines in polychaete jaws (teeth). In *Skeletal Growth of Aquatic Organisms* (Rhoads, D. C. & Lutz, R. A., eds), pp. 561–594. New York: Plenum Press.
- Oliveira, K. (1996). Field validation of annular growth rings in the American eel, *Anguilla rostrata*, using tetracycline-marked otoliths. *Fishery Bulletin, U.S.* **94**, 186–189.
- Pannella, G. (1971). Fish otoliths: daily growth layers and periodical patterns. *Science* **173**, 1124–1127.
- Parrish, B. B. (1958). Comparison of European and North American techniques of measuring nets, of reading ages of fish and of studying growth. *International Commission for the Northwest Atlantic Fisheries Special Publication* **1**, 113–138.
- Pearson, D. E. (1996). Timing of hyaline-zone formation as related to sex, location, and year of capture in otoliths of the widow rockfish, *Sebastes entomelas*. *Fishery Bulletin, U.S.* **94**, 190–197.
- Petterson, G., Odgaard, B. V. & Renberg, I. (1999). Image analysis as a method to quantify sediment components. *Journal of Paleolimnology* **22**, 443–455.
- Planes, S., Laval, P. & Lecomte-Finiger, R. (1991). An otolithic image analysis method for ageing and estimating growth of larvae and juveniles of sea-bass (*Dicentrarchus labrax*): preliminary results. *Scientia Marina* **55**, 591–597.
- Pratt, H. L. Jr & Casey, J. G. (1983). Age and growth of the shortfin mako, *Isurus oxyrinchus*, using four methods. *Canadian Journal of Fisheries and Aquatic Sciences* **40**, 1944–1957.
- Quinn, T. P., Nemeth, R. S. & McIsaac, D. O. (1991). Homing and straying patterns of fall chinook salmon in the Lower Columbia River. *Transactions of the American Fisheries Society* **120**, 150–156.

- Ralston, S. & Miyamoto, G. T. (1983). Analyzing the width of daily otolith increments to age the Hawaiian snapper, *Pristipomoides filamentosus*. *Fishery Bulletin, U.S.* **81**, 523–535.
- Ralston, S. & Williams, H. A. (1989). Numerical integration of daily growth increments: an efficient means of ageing tropical fishes for stock assessment. *Fishery Bulletin, U.S.* **87**, 1–16.
- Rauck, G. (1974). On characteristics of plaice otoliths in the North Sea. *Berichte der Deutschen Wissenschaftlichen Kommission fuer Meeresforschung* **23**, 289–295.
- Richards, L. J., Schnute, J. T., Kronlund, A. R. & Beamish, R. J. (1992). Statistical models for the analysis of ageing error. *Canadian Journal of Fisheries and Aquatic Sciences* **49**, 1801–1815.
- Rittenour, T. M., Brigham-Grette, J. & Mann, M. E. (2000). El Nino-like climate teleconnections in New England during the late Pleistocene. *Science* **288**, 1039–1042.
- Rivard, D. & Foy, M. G. (1987). An analysis of errors in catch projections for Canadian Atlantic fish stocks. *Canadian Journal of Fisheries and Aquatic Sciences* **44**, 967–981.
- Robillard, S. R. & Marsden, J. E. (1996). Comparison of otolith and scale ages for yellow perch from Lake Michigan. *Journal of Great Lakes Research* **22**, 429–435.
- Schmitt, P. D. (1984). Marking growth increments in otoliths of larval and juvenile fish by immersion in tetracycline to examine the rate of increment formation. *Fishery Bulletin, U.S.* **82**, 237–242.
- Schramm, H. L. Jr (1989). Formation of annuli in otoliths of bluegills. *Transactions of the American Fisheries Society* **118**, 546–555.
- Secor, D. H., Dean, J. M. & Campana, S. E. (Eds) (1995a). *Recent Developments in Fish Otolith Research*. Columbia: University of South Carolina Press.
- Secor, D. H., Houde, E. D. & Monteleone, D. M. (1995b). A mark-release experiment on larval striped bass *Morone saxatilis* in a Chesapeake Bay tributary. *ICES Journal of Marine Science* **52**, 87–101.
- Shirvell, C. S. (1981). Validity of fin-ray ageing for brown trout. *Journal of Fish Biology* **18**, 377–383.
- Smith, D. C., Fenton, G. E., Robertson, S. G. & Short, S. A. (1995). Age determination and growth of orange roughy (*Hoplostethus atlanticus*): a comparison of annulus counts with radiometric ageing. *Canadian Journal of Fisheries and Aquatic Sciences* **52**, 391–401.
- Smith, J. N., Nelson, R. & Campana, S. E. (1991). The use of Pb-210/Ra-226 and Th-228/Ra-228 dis-equilibria in the ageing of otoliths of marine fish. In *Radionuclides in the Study of Marine Processes* (Kershaw, P. J. & Woodhead, D. S., eds), pp. 350–359. New York: Elsevier Applied Science.
- Smith, M. K. & Kostlan, E. (1991). Estimates of age and growth of ehu *Etelis carbunculus* in four regions of the Pacific from density of daily increments in otoliths. *Fishery Bulletin, U.S.* **89**, 461–472.
- Spratt, J. D. (1975). Growth rate of the northern anchovy, *Engraulis mordax*, in southern California waters, calculated from otoliths. *California Fish and Game* **61**, 116–126.
- Steffensen, E. (1980). Daily growth increments observed in otoliths from juvenile East Baltic cod. *Dana* **1**, 29–37.
- Stevenson, J. T. & Secor, D. H. (1999). Age determination and growth of Hudson River Atlantic sturgeon, *Acipenser oxyrinchus*. *Fishery Bulletin, U.S.* **98**, 153–166.
- Summerfelt, R. C. & Hall, G. E. (Eds) (1987). *Age and Growth of Fish*. Ames: Iowa State University Press.
- Svedang, H., Wickstrom, H., Reizenstein, M., Holmgren, K. & Florenius, P. (1998). Accuracy and precision in eel age estimation, using otoliths of known and unknown age. *Journal of Fish Biology* **53**, 456–464.
- Thorson, T. B. & Lacy, E. J. Jr (1982). Age, growth rate and longevity of *Carcharhinus leucas* estimated from tagging and vertebral rings. *Copeia* **1982**, 110–116.
- Tsukamoto, K. & Kajihara, T. (1987). Age determination of ayu with otolith. *Nippon Suisan Gakkaishi* **53**, 1985–1997.

- Tyler, A. V., Beamish, R. J. & McFarlane, G. A. (1989). Implications of age determination errors to yield estimates. In *Effects of Ocean Variability on Recruitment and an Evaluation of Parameters used in Stock Assessment Models* (Beamish, R. J. & McFarlane, G. A., eds), pp. 27–35. Canadian Special Publication of Fisheries and Aquatic Sciences **108**.
- Uchiyama, J. H. & Struhsaker, P. (1981). Age and growth of skipjack tuna, *Katsuwonus pelamis*, and yellowfin tuna, *Thunnus albacares*, as indicated by daily growth increments of sagittae. *Fishery Bulletin, U.S.* **79**, 151–162.
- Ueno, S., Imai, C. & Mitsutani, A. (1995). Fine growth rings found in statolith of a cubomedusa *Carybdea rastoni*. *Journal of Plankton Research* **17**, 1381–1384.
- Uozumi, Y. & Ohara, H. (1993). Age and growth of *Nototodarus sloanii* (Cephalopoda: Oegopsida) based on daily increment counts in statoliths. *Nippon Suisan Gakkaishi* **59**, 1469–1477.
- van den Broek, W. L. F. (1983). Ageing deepwater fish species: report of a visit to the United Kingdom September–November 1982. *Miscellaneous Series Fisheries Research Division, Ministry of Agriculture and Fisheries* (unpublished report, MAF Fisheries, Greta Pt Library, Wellington, NZ).
- van der Walt, B. A. & Beckley, L. E. (1997). Age and growth of *Sarpa salpa* (Pisces: Sparidae) off the east coast of South Africa. *Fisheries Research* **31**, 241–248.
- Vilizzi, L. & Walker, K. F. (1999). Age and growth of the common carp, *Cyprinus carpio*, in the River Murray, Australia: validation, consistency of age interpretation, and growth models. *Environmental Biology of Fishes* **54**, 77–106.
- Volk, E. C., Schroder, S. L. & Grimm, J. J. (1999). Otolith thermal marking. *Fisheries Research* **43**, 205–219.
- Waldron, M. E. (1994). Validation of annuli of the South African anchovy, *Engraulis capensis*, using daily otolith growth increments. *ICES Journal of Marine Science* **51**, 233–234.
- Wild, A. & Foreman, T. J. (1980). The relationship between otolith increments and time for yellowfin and skipjack tuna marked with tetracycline. *Inter-American Tropical Tuna Commission* **17**, 509–560.
- Wild, A., Wexler, J. B. & Foreman, T. J. (1995). Extended studies of increment deposition rates in otoliths of yellowfin and skipjack tunas. *Bulletin of Marine Science* **57**, 555–562.
- Wilson, D. T. & McCormick, M. I. (1997). Spatial and temporal validation of settlement marks in the otoliths of tropical reef fishes. *Marine Ecology Progress Series* **153**, 259–271.
- Wilson, R. R. (1988). Analysis of growth zones and microstructure in otoliths of two macrourids from the North Pacific abyss. *Environmental Biology of Fishes* **21**, 251–261.
- Wintner, S. P. & Cliff, G. (1999). Age and growth determination of the white shark, *Carcharodon carcharias*, from the east coast of South Africa. *Fishery Bulletin, U.S.* **97**, 153–169.
- Zhang, Z. & Runham, N. W. (1992). Temporal deposition of incremental and discontinuous zones in the otoliths of *Oreochromis niloticus*. *Journal of Fish Biology* **40**, 333–339.
- Zuk, M. (1987). Age determination of adult field crickets: methodology and field applications. *Canadian Journal of Zoology* **65**, 1564–1566.