EFAN Report 3-2000

Guidelines and Tools for Age Reading Comparisons

Version 1   October 2000

by

A.T.G.W.Eltink, A.W. Newton, C. Morgado, M.T.G. Santamaria and J. Modin

You are kindly requested to refer to this document as:

GUIDELINES AND TOOLS FOR AGE READING. *Eltink, A.T.G.W., A.W. Newton, C. Morgado, M.T.G. Santamaria and J. Modin, 2000. Guidelines and tools for age Reading. (PDF document version 1.0 October 2000) Internet: http://www.efan.no*

And to the workbook as:

Eltink, A.T.G.W. 2000. Age reading comparisons. (MS Excel workbook version 1.0 October 2000) Internet: http://www.efan.no

European Fish Ageing Network (EFAN)

For further information, please contact:
EFAN, Institute of Marine Research, Flødevigen Marine Research Station, N-4817 His, Norway.
Phone: (47) 37 05 90 00; Fax: (47) 37 05 90 01; Email: bente.lundin@imr.no
Office address: Flødevigvn. 49, Hisøy (Arendal), Norway.
Coordinator: Erlend Moksness, Phone (direct): (47) 37 05 90 41; E-mail: moksness@imr.no

GUIDELINES AND TOOLS FOR AGE READING COMPARISONS
Version 1.0   October 2000

A.T.G.W. Eltink, A.W. Newton, C. Morgado, M.T.G. Santamaria and J. Modin

Guus Eltink, Netherlands Institute for Fisheries Research (RIVO),
P.O.Box 68, 1970 AB IJmuiden, the Netherlands
e-mail address: a.t.g.w.eltink@rivo.wag-ur.nl
Direct phone: +31 255 564691

Andrew Newton, Marine Laboratory,
P.O. Box 101, Victoria Road, Aberdeen AB11 9DB, United Kingdom
e-mail address: newtonaw@marlab.ac.uk
Direct phone: +44 1224 295396

Cristina Morgado, Instituto Português de Investigação Marítima,
Avenida de Brasília, 1449-006 Lisboa, Portugal
e-mail address: cmorgado@ipimar.pt
Phone: +351 21 3027061

Maria Teresa Garcia Santamaria,
Centro Oceanografico de Canarias
Ctra. San Andres s/n
38120 Santa Cruz de Tenerife
Islas Canarias
e-mail address: mtgs@ieo.rcanaria.es
Phone: +34 922 549400

Johan Modin, Kristineberg Marine Research Station
Goteborg University
S-45034 Fiskebackskil, Sweden
e-mail address: j.modin@kmf.gu.se
Phone: +46 523 18597

Further copies or updated versions of this paper and the associated MS EXCEL file (Age Comparisons.xls) can be downloaded from the EFAN website:http:/www.efan.no (Activate the Guidelines button).

CONTENTS

1 Introduction

In 1997 the European Commission funded a four year concerted action called the European Fish Ageing Network (EFAN). The stated objectives of EFAN were:

'to develop, conduct and co-ordinate collaborative research and training, and thereby ensure that age determination becomes a reliable element of the assessments underlying the scientific management advice on fisheries and environmental resources.'

Since its inception the network has grown until now there are approximately 50 institutes involved in exchanges of ideas, information and samples. As the network evolved it became apparent that a certain degree of standardisation was required if exchanges of age determination material were to be made efficiently. In addition it was discovered that many individual institutes had techniques and/or protocols which deserved a wider audience than currently experienced. This document is an attempt to formalise and advertise the best features currently available in Europe; the authors have drawn heavily on their own experiences but readers are advised that this is an evolving field and new techniques are constantly being developed.

The guidelines concentrate on:

a) Reference collections and the analysis of the age determinations
b) Running exchange schemes and the analysis of the age determinations
c) Organising workshops and the analysis of the age determinations
d) Digital imaging tools
e) Definition of terms

The essential companion to this document is the workbook **AGE COMPARISONS.XLS** by Guus Eltink, RIVO, IJmuiden, the Netherlands. This is an MS Excel file (Version 1.0, October 2000) which is available on the EFAN website (http://www.efan.no). This workbook is an indispensable tool for the analysis of age determination results by individual readers and is now regarded as the standard method of analysing precision, accuracy etc. The workbook is a self-contained tool with extensive notes, especially on data analysis. Because this document and the workbook are interlinked there are frequent references to the latter in this document. The workbook has two sheets – the first sheet (READMEFIRST) contains instructions and selected outputs are reproduced in Appendix 1. The second sheet (ANALYSIS) contains an input section and a variety of outputs. If you are reading a published copy of this report Appendix 1 also contains examples of all the tables and figures. If you have acquired this report from the internet, examples are **not** shown; in this situation you must view the actual Excel file which is downloaded with this text. In the following text figures and tables from the workbook are referred to in italics e.g. *(Worksheet Figure 3)*.

Within this document frequent reference is made to the co-ordinator or chairperson using the masculine pronoun. This is merely for ease of typing and reading. The authors are well aware of many excellent schemes conducted by members of the fair sex and hope that they, and future co-ordinators/chairpersons, will not be offended or deterred by our editorial decision.

1.1 Why is age reading so important?

Growth (in length and weight), mortality (both fishing and natural mortality), maturity, etc. are time related processes and these are the key items in the analytical assessments of fish stocks. Due to the age reading of the calcified structures these processes can be related to time e.g. weights at age, fishing mortality by year, maturity at age, etc. Errors in age reading will especially affect the estimates of recruitment but also, to a lesser extent, the estimated fishing mortality and spawning stock biomass.

1.2 Age reading is an art, which ought to be checked!

The scientists involved in stock assessment and environmental monitoring are the users of age reading results from the age readers in the different age reading laboratories. The experts should be aware of the quality of these very basic age data because managers often use their biological advice to set Total Allowable Catches (TAC's).

Because age determination is not an exact science there is understandably a 'drift' between readers and within readers over time. Some species present more problems than others. Each age reading laboratory should have in place protocols to ensure that all readers within an institute use an agreed standard age determination method. Within and between age reading laboratories there will inevitably be disagreements between age readers. Within age reading laboratories re-reading control collections at regular time intervals can help to ensure consistency between readers and over time, while agreed and validated collections assist age readers to calibrate their age reading method. A comprehensive exchange scheme of calcified structures between age readers of different age reading laboratories will help to identify differences in age reading. Such an exchange program may identify a problem so serious that a workshop should be held.

This age reading comparison package provides guidelines on how to organise exchanges and workshops and how to analyse data obtained. Furthermore, it provides guidelines for the creation and analysis of reference collections. These guidelines are based on the experience of recent exchanges and workshops and may help the structuring of similar events in the future. The guidelines for the preparation of digitised images are meant to help standardise the production of these images.

This document is designed for co-ordinators of exchanges, chairpersons of workshops and for individuals who analyse reference collections. It is assumed that many of our colleagues may not have much experience on these issues and therefore the guidelines are extensive; they range from basic logistics to sophisticated statistical packages. The main objective of these guidelines is to ensure that the maximum output is obtained from these exercises and that the minimum of work is expended in the analysis of the age reading comparisons itself. Furthermore, these guidelines ensure that age reading comparisons are carried out in a more standardised way, which will enable stock assessment experts to estimate the effect of age reading errors on stock assessments.

1.3 Age reading errors

In principle the age reading errors that affect precision and accuracy have to be estimated. The age reading errors that affect precision are easy to estimate, but those that affect accuracy are in most cases very difficult to estimate, because of the absence of calcified structures of known age.

The term precision is defined as the variability in the age readings. It can be estimated for each calcified structure over all age readings, for each age group by age reader and for all age groups by reader and readers combined. The precision errors in age readings are best described by the coefficient of variation (CV) by age group (CV = st. dev/mean age recorded). A low CV indicates a high precision and a high CV indicates a low precision in age reading (CV is actually more a measure of imprecision). This measure of precision is independent of the closeness to the true age. This CV should be calculated by age group, since it might differ by age group (high CV's for certain age groups might indicate difficulties in age reading these ages e.g. during the period when false rings occur in the juvenile phase!).

The term accuracy is defined as the closeness of a measured value to its true value.

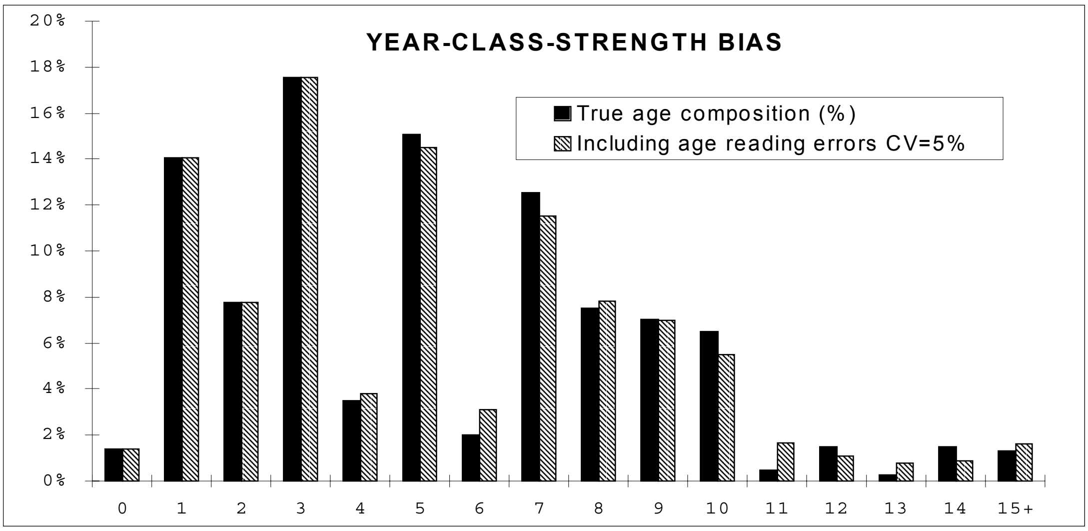
Absolute bias can be defined as a systematical over- or underestimation of age compared to the true age. The accuracy errors are best described by the absolute bias by age group. However, absolute bias can only be measured, if calcified structures of true or known age are available (e.g. from tagging or tank experiments). In the absence of calcified structures of known age, the age readings can be compared to modal age, which is defined as the age determined for an individual structure for which most of the readers have a preference (see also section 1.4).

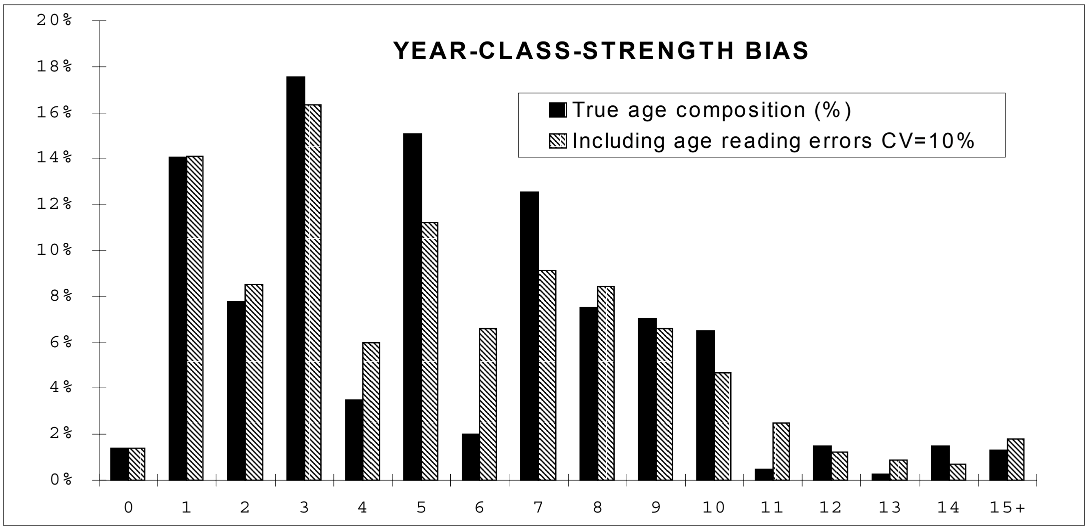
Relative bias can be defined as a systematical over- or underestimation of age compared to the modal age. Relative bias is equal to absolute bias, if there is proof that there is no absolute bias in age reading (based on age reading of even a small set of calcified structures of known age). Only in this special case of no absolute bias in age reading is the estimated mean age similar to the true age.

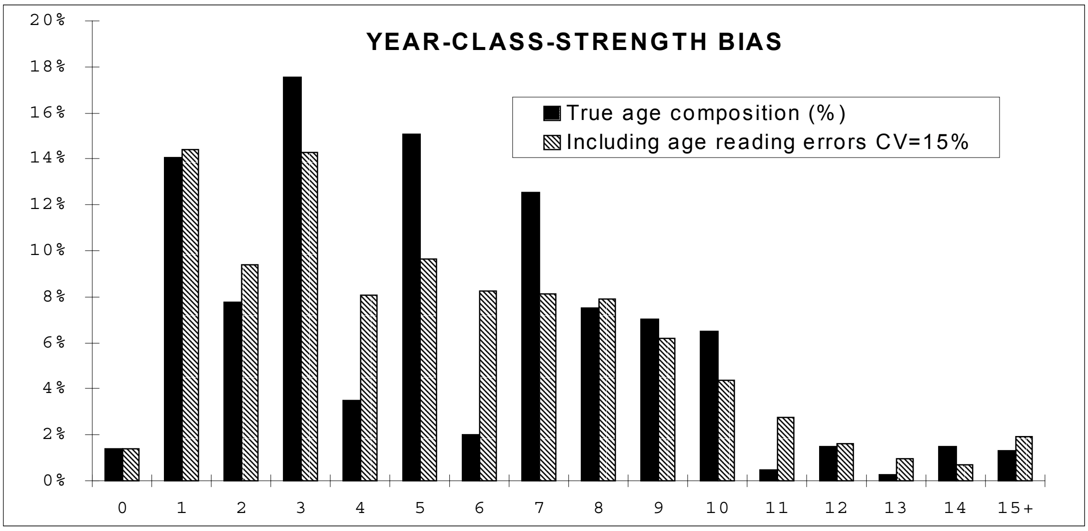
The age reading comparisons to modal age provide a low estimate of relative bias compared to absolute bias, when most readers have a similar serious bias in age reading. The results from a comparison to modal age could be misleading in such case. This problem can only be solved by the use of a set of calcified structures of known ages!

If the age readings of only two age readers are compared, the difference between the readings can be defined as between-reader-bias.

Further errors can be generated due to the relative strengths of neighbouring year classes. In an age distribution sample neighbouring year classes can differ considerably in strength. Due to precision errors certain proportions of fish are assigned to the adjacent younger and older year classes. This causes a bias, which results in an overestimation of the numbers at age of weaker year classes and an underestimation of the stronger year classes (smoothing effect). This year-class-strength bias is defined as the effect of precision errors in age reading on the estimated age composition due to differences in year class strength. This effect of year-class-strength bias is shown in the figures below for three different precision error levels (CV=5%, CV=10% and CV=15%):





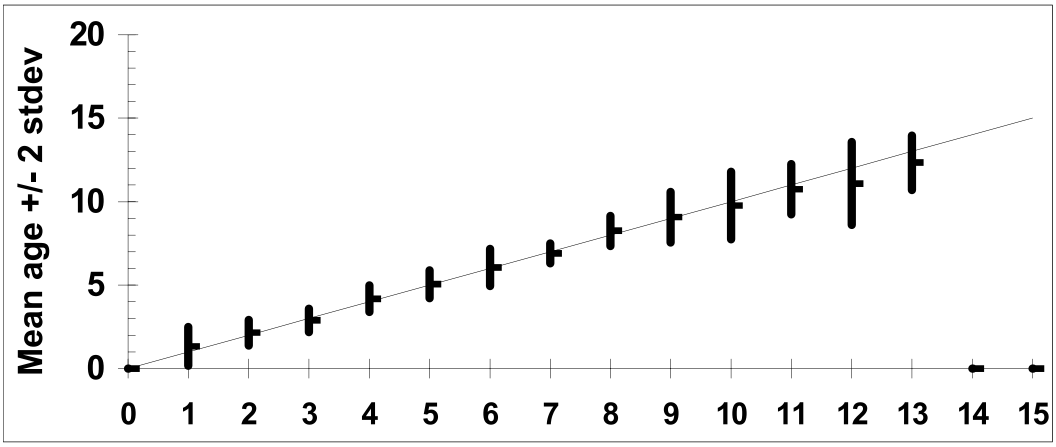


Year class strength bias will increase with higher precision errors (higher CV) but it should be emphasised that this type of bias is not directly caused by age reading errors in accuracy, but only indirectly by precision errors in age reading. It is especially noticeable if the numbers of calcified structures read at each age are considerable different. However, this year-class-strength bias is relatively small, if precision errors are low (CV <5%) and/or if adjacent age groups are approximately equally represented in the age sample.

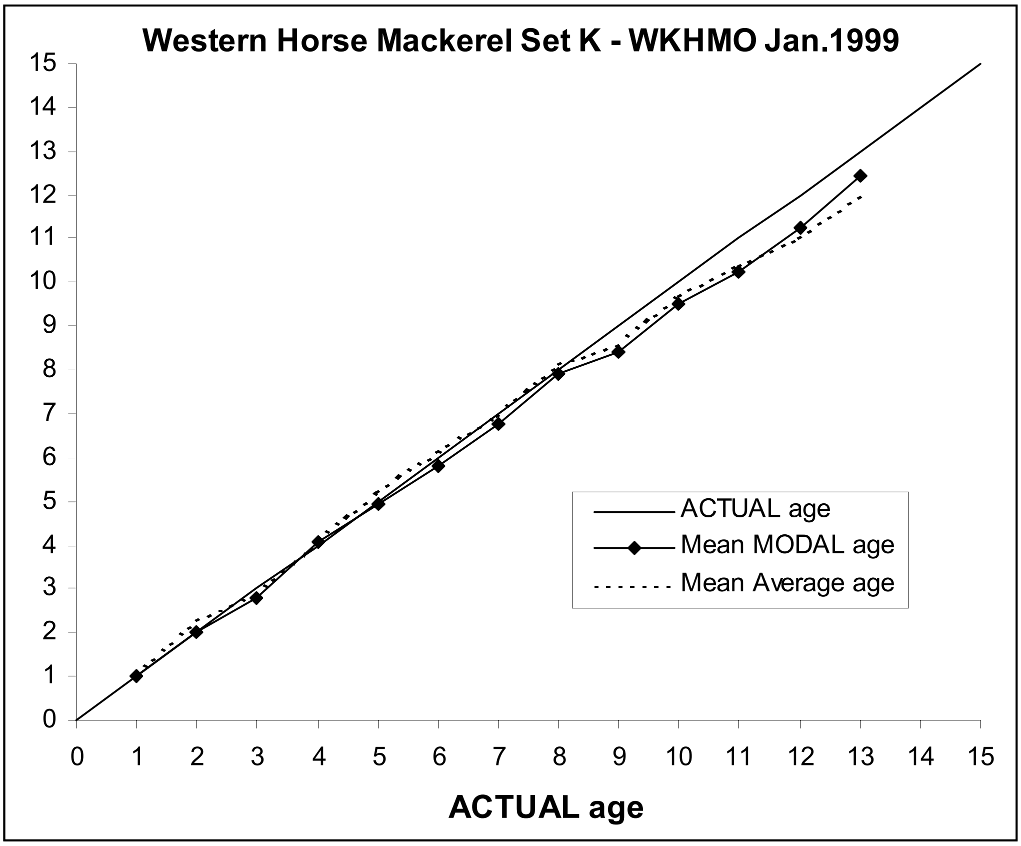
Finally, there exists yet another error in age reading, which is called "reader bias", but which should not be estimated in exchanges or at workshops. This "reader bias" is defined as the subjective assignment of ages, which is caused by the knowledge of existing strong or weak year classes. This causes strong year classes to become stronger and weak year classes to become weaker. Normally the age reading errors that affect precision will reduce the strength of strong year classes and increase the strength of weak year classes (year-class-strength bias). However, the reader bias has an opposite effect, because the age reader tries to reduce the spread in age readings by taking into account his knowledge about year class strength! This "reader bias" is an age reading error, which should not occur in the age reading comparisons (unless it is an objective to estimate it). Furthermore the knowledge of length can also cause a reader bias e.g. relatively small fish could be assigned to lower ages and relatively large fish to older ages.

1.4 Age bias plots and comparisons to modal age

In age reading comparisons it is recommended to make age bias plots in which the mean age recorded ±2stdev is plotted against the modal age (ICES, 1994 and Campana *et al.*, 1995). Such figures are extremely useful in presenting the differences in the age reading methods between two readers.



At the 1995 Workshop on Mackerel Age Reading (ICES, 1995) examples were given of age bias plots. This way of comparison had the advantage that the age reading methods of the individual age readers were compared to modal age, assuming that the modal age was the best and closest approximation to true age. This assumption was tested with mackerel otoliths of known age (tagging experiments), which were available at that workshop (ICES, 1995). There appeared to be a good linear relationship between known age and modal age, which indicated that the estimated modal age appears to be a good approximation to known or true age (average age underestimated the older ages). Therefore, in the absence of calcified structures of known or agreed age, the age readings in the *Worksheet* are compared as a standard to the modal age. However, it is recommended that this basic assumption of the modal age being the best approximation to the true age be tested whenever validated calcified structures are available (see example below for horse mackerel).



2 Reference collections

2.1 Introduction

There appears to be some confusion concerning the term reference collection as it seems to have several interpretations by fishery scientists. The authors contend that 'Reference Collections' should be merely regarded as a generic term to include several types of completely different types of collections. In the following text three types of collections are distinguished and they cover sets of calcified structures of fishes whose ages are either unknown or agreed or known .

2.2 Collection of unknown ages (Control collection)

Control collections are sets of calcified structures of unknown age, which are re-aged at regular time intervals by experienced age readers to estimate changes in the precision and relative bias at age over time. These control collections should contain calcified structures, which are representative of the population at that point in time (i.e. they have not been selected as good quality calcified structures!). The estimated precision and relative bias can be assumed to be representative for the age readings, which the age reader provides for stock assessment purposes. Changes in precision and relative bias can be observed over time, if the control collection is re-aged at regular time intervals (e.g. each year) and the age comparisons are related to modal age. Therefore, control collections can prove that the age reading method of an age reader has been consistent over a long time period. Furthermore it can prove that the age reading method of replacing age readers is consistent as well. It is important in the evaluation of stock assessment problems to be able to exclude inconsistent age determination as a factor. Furthermore these control sets can be used as training tools for new age readers. However, agreed and validated collections are more appropriate for this purpose (see sections 2.3 and 2.4). Whichever type of collection is used should be large enough to prevent immediate previous image recognition to the readers.

Each age reading laboratory should have control collections for each species/stock/season but it should be borne in mind that, over time, environmental factors may affect the micro and macro structures displayed by calcified structures. Thus any such collection might be updated at frequent intervals. The analyses of age readings of a control collection should represent a history book on the precision and relative bias in age reading regarding all age readers within an age reading laboratory. It should show whether significant changes in precision and relative bias took place over time. Such control collections might also be used to test whether the precision in age reading of an age reader is good enough before he starts age reading after a long period of not reading that particular species. Furthermore such control collections might be used for quality assurance, if a threshold is set on the precision in age reading before an age reader can become a qualified age reader for a certain species.

The criteria for interpreting calcified structures are transferred from person to person. With a control collection acting as a reference, subtle shifts in age reading between age readers can be detected. Past interpretations that are found incorrect in the light of new data can be found and corrected. The use of control collections facilitates the management of age reading teams, because it can provide early signs of divergence and differences in perception. An evaluation, that is perceived to be fair, is more likely to elicit positive adjustments between readers (ICES, 1994).

The preparation technique of the calcified structures for control collections is extremely important because these collections should have a high durability (e.g. sliced or whole otoliths embedded in resin under glass cover slips). It is not acceptable that the quality of the individual calcified structures deteriorates during the course of constant re-reading because it will reduce the precision in age reading over time. In order to avoid this deterioration the control collection might evolve slowly with the periodic addition of new structures and by the deletion of damaged structures having approximately the same age. The mean CV can not be compared over the whole time period, if the age composition is changed during the course of time. If the age composition is changed only CV at age and relative bias at age can be compared over the whole time period.

Mean CV's at age are expected to be comparable between control sets of different age reading laboratories if the preparation techniques of the calcified structures and the readability of calcified structures from different areas are approximately the same. But mean CV's averaged over all calcified structures are not expected to be comparable, because separate control sets are likely to consist of different age compositions (assuming that CV is not equal for all age groups).

2.3 Collection of agreed ages (Agreed collection)

Agreed collections are sets of calcified structures of unknown age which have achieved at least an 80% agreement when aged during an exchange or at a workshop. It can be assumed that the calcified structures selected for the agreed collection are easier to age than the usually random selected calcified structures as used for age reading for stock assessment purposes. Therefore, these agreed collections are not suitable for estimating a representative precision and a relative bias i.e. an Agreed Collection cannot perform the function of a Control Collection. These agreed collections are especially useful as training tools for new readers and for calibrating the age reading method of both experienced and inexperienced readers if calcified structures of known age are not available.

One of the problems with agreed collections is that they might be composed of selected calcified structures from several exchanges or workshops and can be the property of different age reading laboratories. The age readers involved in age reading this particular species will often have difficulties in obtaining and reading an agreed collection. Therefore, the production of digitised images of an Agreed Collection is recommended as this is a very good alternative to making the original collection available to all age readers (see section 5.7).

2.4 Collection of known ages (Validated collection)

Validated collections are sets of calcified structures of known age (either direct validation e.g. from tagging or tank experiments or microstructure research or indirect validation e.g. from extremely strong year classes or otoliths from the first strong modes of the length distribution). These validated sets are especially useful to calibrate the age reading method of both experienced and inexperienced readers. Furthermore they can be used as training tools for new readers. It is recommended to produce digitised images of these validated collections, which is the best way to make them available to all age readers.

For training and calibration purposes validated collections should be preferred above agreed collections because they contain calcified structures of known age. This implies that if enough calcified structures of known age are available only a validated collection has to be made. However, the number of Validated Collections known to exist in Europe is minimal and all institutes are encouraged to source material for such collections.

2.5 Summary

Gradual changes in the precision and the accuracy of age reading of an age reader might occur over a long time period. Furthermore abrupt changes could take place when age readers are replaced. **Control** collections of calcified structures (otoliths, scales, bones, etc.) can be used with the objective of detecting these changes over time (e.g. when these collections are read annually) and to reflect a representative estimate of precision and relative bias in age reading as provided for stock assessment purposes. The age reading comparisons from control collections will detect possible changes in the age reading method over time, if the age reading method has been changed due to calibrating it with agreed or validated collections.

Agreed and **Validated** **collections** can be used with the main objective of calibrating age reading methods of both experienced and inexperienced age readers. Furthermore these are very useful for training new age readers. These agreed and validated collections are also very useful for image analysis experts who, in developing automatic age reading devices, need image data bases of agreed or validated collections to test their age reading algorithms.

2.6 Analysis of age readings from collections

The MS Excel *workbook* "*AGE COMPARISONS.XLS*" has been designed as a standard statistical tool for analysing age readings from calcified structures that have been exchanged or that were obtained at age reading workshops. However, it can be used for the analysis of age reading results from control, agreed or validated collections as well. In such cases only small changes have to be made to the *worksheet "Analysis"*. An explanation on age reading comparisons is given in section 3.13 (Age reading comparisons).

2.6.1 Amending worksheet for Control collection

The title above *Table 1* should refer to the name of the control collection, the species and the name of the age reader. In *Table 1* country and initials of the age readers should be replaced by the dates of age reading or re-reading. If the number of times that the collection is aged exceeds 15 the spreadsheet should be extended (see notes in workbook). Instead of different readers the workbook now refers to the same age reader re-reading the same set several times. The output figures and tables of *sheet "ANALYSIS"* will then show for the same age reader the changes over time in the precision of age reading (CV), in the relative bias and in the percentage agreement.

2.6.2 Amending worksheet for Agreed collection

Age readings of many age readers compared to an agreed age
The following changes to *worksheet "ANALYSIS"* have to be made. The column, where modal age is calculated, should be used to fill in the agreed age (change the heading to "AGREED age"). Everywhere in the spreadsheet where "MODAL age" is mentioned, it should be replaced by "AGREED age". The bias test should be carried out by testing the ages of each reader against the agreed ages.

Re-readings of one age reader compared to an agreed age
The following changes to *sheet "ANALYSIS"* have to be made. The title above *Table 1* should refer to the name of the agreed collection, the species and the name of the age reader. In *Table 1* the country and initials of the age readers should be replaced by the dates of age reading or re-reading. If the number of times that the collection is read exceeds 15 the spreadsheet should be extended (see notes in workbook). The column, where modal age is calculated, should be used to fill in the agreed age (change also the heading to "AGREED age"). Everywhere in the spreadsheet where "MODAL age" is mentioned, it should be replaced by "AGREED age". The bias test should be carried out by testing the ages of each time of reading against the agreed ages.

The bias plots and the tables with CV and relative bias will show the errors in the age reading method of this particular age reader. These will help him in correcting his age reading method. Using defined time intervals he can re-read this agreed collection to check whether he has achieved an improvement. Instead of different readers the spreadsheet refers now to the same age reader re-reading several times the same set. The output figures and tables of *worksheet* "*ANALYSIS*" will now show for the same age reader the changes over time in the precision in age reading (CV), in the relative bias and in the percentage agreement. It should be remembered that the precision and bias are probably not representative for the age readings presented for stock assessment purposes since the age readings are highly likely to have been made only on good quality calcified structures.

2.6.3 Amending worksheet for Validated collection

Age readings of many age readers compared to known age
The following changes to *worksheet "ANALYSIS*" have to be made. The column, where modal age is calculated, should be used to fill in the actual age and the heading

changed to "ACTUAL age". Everywhere in the spreadsheet where "MODAL age" is mentioned, it should be replaced by "ACTUAL age". "RELATIVE bias", it should be replaced by "ABSOLUTE bias" and in *Table 3* the text of "Reader against MODAL age bias test" should be replaced by "Reader against ACTUAL age bias test".

Re-readings of one age reader compared to known age
The following changes to *worksheet* "ANALYSIS" have to be made. The title above *Table 1* should refer to the name of the validation collection, the species and the name of the age reader. In *Table 1* the country and initials of the age readers should be replaced by the dates of reading or re-reading. If the number of times that the collection is aged exceeds 15, the spreadsheet should be extended by the appropriate number of columns. The column, where modal age is calculated, should be used to fill in the actual age and the heading to changed to "ACTUAL age"). Everywhere in the spreadsheet where "MODAL age" is mentioned, it should be replaced by "ACTUAL age". "RELATIVE bias", it should be replaced by "ABSOLUTE bias" and in *Table 3* the text of "Reader against MODAL age bias test" should be replaced by "Reader against ACTUAL age bias test". The bias test should be carried out after each re-reading by testing the obtained ages against the actual ages.

The comments about bias plots etc. made in the last paragraph in section 2.6.2 are equally applicable in this section.

3 Exchanges

3.1 Introduction

The objective of exchanges of calcified structures is to estimate precision (CV) and relative / absolute bias in the age readings from age readers of the different age reading laboratories.

The differences in age reading might not only be caused by different age reading methods, but might also be caused by different preparation methods of the calcified structures (e.g. use of burnt otoliths and whole otoliths) or the use of different calcified structures (e.g. the use of scales and otoliths). Age reading methods of the age readers can only be compared, if all participants of the exchange accept the same method of preparing the calcified structures for age reading and actually use the same calcified structures for age reading.

An exchange program of calcified structures (otoliths, bones, scales, etc.) can be carried out to test whether significant differences in age reading methods exist between age readers and to test whether different calcified structures or preparation methods of calcified structures cause significant differences in age reading results. Priority should be given to an exchange when age readings have to be compared from age readers of different age reading laboratories. The costs of an exchange are much lower than of an age reading workshop, which should only be organised when the problems in age reading have become evident from an exchange. A workshop should then aim to solve these specific problems in an efficient way.

An exchange program should be carried out regularly e.g. every second year to check whether the precision in age reading and bias of the age readers is still within acceptable levels. Exchange programs obtain more objective estimations of the precision and bias in age reading, since the readers use their own equipment and are not subject to a tight time schedule (criteria which may not be applicable in a workshop).

The following points should be considered in establishing an exchange scheme; a summary is made in the flowchart (Appendix 2).

3.2 Identification of possible problems

At the start of an exchange it is very important to have some idea of the problems that might cause differences in age reading between the possible participants in the exchange. Generally these problems divide into categories:

1 The use of different calcified structures or different preparation techniques by age readers (see section 3.7)
2 The application of different age reading methods by the age readers, which might be indicated by the following features:
   - Large differences in growth parameters within the same population;

- The interpretation of the edge of calcified structures usually causes more problems in age reading when the calcified structures are collected in the period of fast growth. This can be tested by comparing the age reading results within one set of calcified structures collected in the period of slow growth and fast growth *(see Worksheet Table 4)*
- The interpretation of the annual rings in calcified structures might be more difficult, because of the occurrence of false rings during the juvenile period. This is indicated by higher CV's for the younger ages compared to those for older ages.

3.3 Requirements for a set of calcified structures

The age span in an exchange set of calcified structures should, if possible, be from age 0 to the maximum age possible (try to exceed the age range as used for stock and environmental assessment purposes). The exchange should provide information on the precision in age reading by age group and all age groups combined for each age reader. The precision in age reading is indicated by the coefficient of variation (CV) at age. As a rule of thumb one could say that at least 10 specimens are needed within each age group for a reliable estimation of CV at age. This implies that when dealing with a species with a long life span e.g. mackerel with 20 age groups, an exchange would need at least 200 otoliths within one otolith set, but for anchovy (4 age groups) only 40 otoliths. Therefore try to select the required number of calcified structures for each age group. Randomly taken fish samples will result in far too high a number of age readings of young age groups and a shortage of age readings for the older age groups.

A much higher number of calcified structures is needed, if CV's by month/stratum and by modal age have to be compared to identify where the problems in age reading may lay, e.g. during the period of opaque material deposition and/or during the juvenile period because of false ring formation (*see Worksheet Tables 4 and 5*). In this case 10 specimens within each age group are recommended for each month or stratum.

It is recommended that calcified structures are included in a set in such way that the number with translucent edges and the number with opaque edges are representative of the annual distribution. This is to ensure that the estimated precision and bias are representative for the age readings over the whole year as used for stock assessment purposes. (*Figure 4* will often show that the precision in age reading is worse during the time of opaque edge formation when compared to the time of translucent edge formation)!

The co-ordinator might also decide to assemble a set of calcified structures, which consists of a number of sub-sets. In *Worksheet Table 1* each sub-set can be given a name for the stratum in column A. The co-ordinator could ask each participant or some of the participants to send him a sub-set of calcified structures to be combined into the one set of the exchange. This exercise might show that there are differences in the precision in age reading by sub-set, which might be due to the preparation technique or which might indicate that otoliths from certain areas are more difficult to read, etc.

Each calcified otolith exchanged should bear a unique identification number. This enables the co-ordinator to identify the source of each otolith and also flag it with essential biological data. Whilst the co-ordinator may not wish to circulate all the available biological data at the onset; at the very minimum they must provide age readers with information on area of catch, fish length and month of catch. If growth is different for males and females, sex should be added as well. The information on month of catch is especially important for the interpretation of the otolith edge to determine whether one has to count the translucent/opaque rings or not!

The calcified structures within a set could be ordered according the length of the fish forcing the age readers to start with the smallest fish in the set and gradually increasing to finish with the largest fish. In this manner the readers are able to understand the underlying growth pattern inherent in each species/stock.

The number of possible age reading problems, that you want to check, determines the number of sets in the exchange. However, the co-ordinator should try to limit the total number of calcified structures to 400 otherwise the burden for the age readers will be too much.

3.4 Extra set of calcified structures with VALIDATED ages

The co-ordinator should inquire whether calcified structures of known age are available to be included as an extra set in the exchange. He should do his very best to include such a separate set of calcified structures of known age, even if there are only a few available. The results of the age reading comparisons to modal age are very much the same as the results of comparisons to true age, if the age reading results from even a small validated set indicate that there is no absolute bias in the age readings! However, it should be taken into account that the results of age reading comparisons to modal age will overestimate percent agreement and underestimate the bias, if a strong absolute bias in age reading occurs.

3.5 Use of digitised images

Digitised images can be made from various thin sliced calcified structures because of their perfectly flat surface (possibly also scales when pressed between glass plates). In general, whole otoliths and transversely broken otoliths/bones do not have such a flat surface and need constant focussing when reading the annual rings. Therefore, these can not, currently, be used to make digitised images for an exchange.

Digitised images, in addition to the calcified structures, can be very helpful when the first year's growth zone (see next section) has to be measured and when the nucleus and the annual rings have to be marked on the images. This is an excellent method of checking how each individual reader arrives at their estimated age because, obviously, each reader assigns the annual rings according to their own interpretation.
Further guidelines on digitised images are given in section 5.

3.6 Use of measurements in calcified structures

Microscopic measurement of ring growth can also be an aid in age estimation of calcified structures. For example, an indirect evidence of the validity of otolith or scale reading can be obtained by comparing the back-calculated total fish length with the known lengths that the fish attained in previous years. In many species the body/otolith length relationship assumes a straight line, thus a direct proportion method can be used to back calculate the body lengths obtained at the end of each annulus (Liew, 1974, Francis, 1995).

In the art of reading otoliths a slightly different measurement can aid the reader in detecting the true annual rings. D_1 is defined as the greatest distance (on any plane) between the outer edges of the first translucent ring; this corresponds to the first year's growth zone in the otolith. Similarly $D_2 – D_n$ are defined as the greatest distances between the outer edges of any subsequent translucent rings. The measurements of D_n are not taken from the nucleus to the outer edges of the translucent rings, because the exact point of the nucleus is often not well defined. The general rule is that D_2-D_1 should be greater than D_3-D_2; D_3-D_2 should be greater than D_4-D_3, etc. In other words: the growth increment at age 1 should be greater than the growth increment at age 2 and the growth increment at age 2 should be greater than growth increment at age 3, etc. Measurements that fail to obey this general rule will indicate the possible occurrence of false rings wrongly counted as annuli. A co-ordinator might request the participants in the exchange to undertake some measurements, e.g. D_1, D_2, D_3 and D_4 in an agreed unit of measurement (e.g. µm) to ensure comparability. This method enables him to discover those calcified structures in the exchange set which cause difficulties in the interpretation of the inner rings. Furthermore it assists him to identify the readers, who often count false rings as annual rings. But the most relevant point is that this method guides the age readers to take a sound decision when evaluating the inner rings.

3.7 Comparison of different preparation techniques or calcified structures

It is recommended that a co-ordinator of an exchange first tries to make an inventory of the calcified structures that are used for age reading and an inventory of the different preparation techniques that exist among the age readers, who will participate in the exchange. The co-ordinator should inquire whether the age readers have tried to improve the preparation method. If so, this can be tested in an exchange.

3.8 Comparison of preparation techniques by age reading laboratory

The co-ordinator should ask the participants of the exchange to send to him a set of calcified structures, which have been prepared in the standard way of the age reading laboratory (together with information on area of catch, month of catch, length, sex, sample year, sample number, fish number). These small sets by laboratory can be merged to one exchange set. In the analysis *(see Worksheet)* the precision, agreement and bias is calculated for each sub-set separately (by stratum), which enables an evaluation of the different preparation techniques by age reading laboratory based on the age readings of all age readers. However, because a particular preparation method causes age reading problems to most readers it does not necessarily mean that the age reader of that particular age reading laboratory has problems in age reading (he might be accustomed to his own preparation technique, but other readers are not).

3.9 Comparison of sets of different preparation techniques or of different calcified structures

If one of the objectives of the exchange is to compare different preparation techniques then two calcified structures should be taken from the **same** fish e.g. both *sagitta* otoliths should be extracted. A comparison of different preparation techniques can then be carried out e.g. whole otoliths compared to broken/burnt transverse sections of otoliths or broken/burnt transverse sections of otoliths compared to thin sliced transverse sections of otoliths etc. Similarly one set of calcified structures can be compared to a second set of another calcified structure if two sets of calcified structures e.g. otoliths and scales are taken from the **same** fish. This technique reduces variance associated with the usage of different fish.

3.10 Participation

The co-ordinator should contact other age reading laboratories and discover whether age readers are interested in participation in an exchange. At the same time he should also inquire how much experience the readers have in age reading this species and other species. The co-ordinator should make a list of participants for the exchange. If the age readings are to be compared to modal age, the co-ordinator should be aware of the fact that modal age can not be estimated if all age readers provide a different age (see notes in *workbook* or Appendix 1). Therefore, the co-ordinator should try to adjust the number of participating readers to achieve a reliable modal age (previous experience suggests that 5 readers is sufficient in most cases). For a species with a small number of age groups fewer age readers are required for the calculation of modal age (e.g. anchovy with 4 age groups) than for a species with a large number of age groups (e.g. mackerel with 20 age groups). It is recommended to try to carry out an age reading comparison to modal age, because it provides information on the precision in age reading by age group for each individual age reader and it provides information on the relative bias. The age bias plots are especially helpful for age readers to calibrate their age reading methods.

Only between-reader-bias can be estimated, if a small number (2 or 3) of age readers participate in the exchange (this is too low a number of age readers to estimate modal age). In a comparison to actual or agreed age it is sufficient to have only one age reader, because modal age does not have to be estimated.

3.11 Instructions from co-ordinator to participants of the exchange

The co-ordinator of the exchange should explain in his letter to the participants the purpose of this exchange and why a certain number of sets have been selected (background problems). He should ask the participating age readers to complete the input forms for age reading results and inform them that the essential information on month of catch, area of catch and length of the fish (and sex if growth is sex dependent) is provided on this input form. All age readers should get one input form for each set of calcified structures for entering their age reading results. The age readers have to enter one estimate of age per calcified structure, which then should be the best estimate. The programs for age reading comparisons can only handle one age estimate and not uncertain age estimates such as “6 or 7”. They should not use a + group (combining fish of the older ages in one group e.g. 10 years and older). The age readers should not omit any calcified structures for age reading, because it influences the results on precision and accuracy (the age readers are only comparable, if they all read the easy and difficulty calcified structures). If the age readers have comments on certain calcified structures they should annotate the input form, referring to the identification number of the calcified structure. This is especially important when calcified structures become damaged during the exchange or when these are put in the wrong envelopes, etc. The co-ordinator can exclude these from the final comparison of age reading results.

The co-ordinator should request the participants to be extremely careful with the sets of calcified structures, since damage to the material will reduce the precision in age reading for subsequent age readers!

The co-ordinator should give a short description of the age reading criteria that should be adopted when age reading: e.g. date of birth assumed to be 1st of January, etc. Furthermore he should give guidelines how the calcified structures have to be viewed: e.g. with reflected or transmitted light, with or without black background, etc. The co-ordinator should send a contact list to all participants of the exchange. This list should not only contain the names but also the full addresses, phone/fax numbers and e-mail addresses. Post box addresses should be avoided because if special couriers are used deliveries cannot be made to such addresses.

The co-ordinator should provide references or send copies of reports of earlier validations, exchanges and workshops concerning this species.

The co-ordinator should mention in his communication that the sets will be sent separately and that he will inquire when each reader has time available to carry out the age readings.

3.12 Circulation

One of the major problems in an exchange of calcified structures is the length of time taken for the successful completion of an exchange scheme. It has to be acknowledged that exchange schemes have a reduced priority compared to an institute’s own work. It is incumbent on the co-ordinator to organise the exchange as efficiently as possible by being pro-active. There is an apocryphal story of one such exchange lasting 10 years!

It is recommended that all age readers read the same calcified structures in the exchange in order to maintain the direct comparability of the age reading results. If there are a large amount of calcified structures to be aged then consideration should be given to dividing the material into sets of approximately 150 - 200 calcified structures. (This is to avoid a heavy burden when sending all calcified structures at the same time). At the start of the exchange these sets should be sent to different participants. Conversely, if there are a large number of participants the material could be divided into two sets and again circulated in two different directions. However, the co-ordinator must ensure that all participants will, eventually, read all the calcified structures (Appendix 3).

The co-ordinator should keep track of where the sets are, who is reading them and to whom each set should be sent next. The co-ordinator should recommend sending the sets by special courier (see note in 3.11) in order to speed up the exchange and to reduce the possibility of losing one of the sets.

The sets should be very well wrapped, in order to avoid damage during transport. An age reader should contact the co-ordinator of the exchange as soon as he has finished his age reading and he should send the co-ordinator (not the other participants!) his age reading results on the input form. The co-ordinator should inquire who of the other age readers has time available for reading that particular set and should inform the age reader to whom he has to send this set. In this way the co-ordinator always knows where the sets are and how long age readers take to carry out the age readings.

If digitised images are available for an exchange they should be circulated for annotation purposes in order to compare age reading criteria of the different participants. Technical advances are very rapid in this field and the co-ordinator should investigate the most efficient method of circulating these images e.g. remote access, posting on the web etc.

3.13 Age reading comparisons

There are several ways of comparing age readings. However, the best way is by making age bias plots, which are easy to understand for the age readers (ICES, 1994 and Campana *et al*., 1995). They enable them to improve their age reading method. They can compare their own age reading method with those of other readers and the whole group. However, it is not always possible to make these age bias plots, because this can depend on the number of participants in the exchange and whether validated or agreed calcified structures are available or not. Therefore, the explanations on how to compare the age readings are divided in the following three sub-sections (references – *in italics* – refer to either Appendix 1 or the *Workbook* depending on the text version):

3.13.1 No age validated material available

Age bias plots are a perfect way of showing the age readers both types of age reading errors (affecting precision and accuracy), when calcified structures of known or actual age are available (see also section 1.3 on age reading errors). However, if no age validated calcified structures are available, then only the age reading errors that affect precision can be estimated. In this case the bias in age reading (accuracy error) can only be shown as relative bias.

Table 1 contains the sample information and the input data of the age readings by reader. It is important to note that the calcified structures, which appear to be difficult for age reading, are indicated with a high CV and a low percent agreement. At workshops age readers should discuss these difficult calcified structures in order to get an agreement on their interpretation.

Figure 1 shows the age bias plots in which the mean age recorded ± 2stdev is plotted against the modal age. The relative bias, which is an age reading error that affects accuracy, corresponds to the difference between mean age recorded and modal age. The age readings are in agreement with modal age when the mean age recorded is on the 1:1 equilibrium line (mean age recorded equal to modal age). Readers have a relative bias in age reading when the mean age recorded is lower (underestimation of age) or higher (overestimation of age) than the modal age. The age reading errors affecting accuracy are best described by the relative bias by age group for each age reader, if calcified structures of known age are not available. However, it should be taken into account that relative bias might provide a very serious underestimate of absolute bias, because the comparison is not made to known/actual ages of the calcified structures!

The precision errors are indicated by ± 2stdev in *Figure 1*. The lengths of the error bars indicate the spread in the age readings. A high precision is achieved, if the error bars remain relatively small. However, in *Tables 1, 2, 4 and 5* it is preferred to show the precision errors not as standard deviation, but as coefficient of variation (CV = STDEV/mean age estimated), because the standard deviations increase greatly with age, while CV remains far more stable, since it is much less age dependent *(see Figure 2).* Relatively high CV's for certain age groups indicate specific problems in age reading. The precision errors by age reader are best described by the coefficient of variation (CV) by age group, because the CV might often be different by age group.

Table 2 shows the precision (CV) in age reading by modal age, by age reader and all readers combined. Relatively higher CV's for the younger modal age groups are expected, when age reading difficulties occur, e.g. because of false rings during the juvenile period. The weighted mean CV's by age reader indicate the relative precision in age reading by reader. These can only be used to compare the precision levels in age reading of the age readers for that particular set (possibly age dependent!). The age readers are ranked according the precision they achieved.

Table 2 shows for each reader the weighted mean percentage agreement to the modal age over age groups 0-15. These mean agreements are related to accuracy, but should not be used to express accuracy reached by each reader, because they are very age dependent *(see Figure 2)*. Percentage agreement decreases significantly when modal ages increase. Percentage agreements are, therefore, only representative by age group. The readers are ranked according to their achievement in the weighted mean percent agreement. It should be preferred to express the accuracy reached by reader as relative bias by age group. Relative bias by age group can be assumed to be equal to absolute bias by age group, if from another set with validated age structures it can be proven that there is no absolute bias in age reading! A rough indication of the relative bias by reader is the weighted mean of the relative bias over all age groups by age reader. The age readers are ranked from lowest to highest relative bias. *Figure 4* shows the relative bias by modal age as estimated by all age readers combined.

At the end of *Table 2* the age readers are ranked based on the average ranking according to CV, agreement and relative bias.

Figure 3 shows the distribution of the age reading errors in percentage by modal age as observed for the whole group of age readers. The achieved precision is shown by the spread of age reading errors. The distributions are skewed, if relative bias occurs.

Figure 1 and *Table 2* present the age reading method of each individual age reader assuming that modal age represents the best age available for comparison. It shows how the age readers have to correct for the bias in age reading (assuming modal age is correct!) and for what ages they have to try to improve the precision (i.e. reduce the CV). At a workshop *Figure 1* and *Table 2* are the most important tools to show age readers how they have to improve the age reading method.

Table 3 shows the age compositions and the mean length at age obtained by each reader and all readers combined. The results on mean length at age by reader are best presented in *Figure 5*, which might illustrate age reading problems especially in the younger fish of certain age readers.

The minimal requirement for age reading consistency is the absence of bias among readers and through time. The hypothesis of an absence of bias between two readers or between a reader and the modal age estimates can be tested non-parametrically with a one-sample Wilcoxon signed rank test. A work table for the calculation of a between-reader bias test is included on the right hand side of *Table 1*. Age reading results of only two readers can be copied from *Table 1* and pasted in the two columns of the work table. The result of each test has to be entered manually in each cell of the tables "Inter-reader bias test" and "Reader against modal age bias test" (see *Worksheet*).

Table 4 shows the precision, percentage agreement and bias achieved by month by modal age. Higher CV's might be observed during the period of opaque deposition on the edge of the calcified structure, because this often causes difficulties in the interpretation. This *Table 4* might also show difficulties in the interpretation of false rings during the juvenile period (for the younger modal age groups), but now by month instead of by reader as in *Table 2*.

Table 5 shows the precision, percentage agreement and bias achieved by stratum. The strata could be sub-sets of calcified structures by age reading laboratory or calcified structures from different areas or calcified structures from different preparation techniques or combinations of months etc. Higher CV's for a certain stratum might indicate that these calcified structures are more difficult to read for some reason.

Table 6 shows the actual values used for producing the age bias plots of *Figure 1*.

Table 7 shows the number of calcified structures by modal age for which at least 80% agreement was obtained by the readers. These calcified structures can be used for the agreed collection (see section 2.3).

3.13.2 Set of age validated material available

Age bias plots are a perfect way of showing the age readers both types of age reading errors (precision and accuracy errors), especially when calcified structures of known or actual age are available. Since all age readings are directly compared to the known/actual age, only one age reader is sufficient for a comparison. By including more readers it becomes more evident what levels of precision and accuracy in age reading can be reached by the best age readers and what levels of precision and accuracy should be aimed for in the future (setting of thresholds for quality assurance).

It is recommended that the basic assumption of the modal age being the best approximation to true age be tested whenever validated calcified structures are available (see section 1.4). This can be done by plotting true age against the modal age as estimated from the age readings of the group.

The worksheet is made for age comparisons to modal age. However, this sheet can easily be changed for a comparison against actual age. (See section 2.6.3).

3.13.3 Set of agreed age material available

Since all age readings are directly compared to the agreed age, only one age reader is sufficient for a comparison. By including more readers it becomes more evident what levels of precision in age reading can be reached by the best age readers and what levels of precision should be aimed for in the future (setting of thresholds for quality assurance). The accuracy is indicated by relative bias, since the known/actual information is lacking.

The worksheet is made for age comparisons to modal age. However, this sheet can easily be changed for a comparison against agreed age. (See section 2.6.2).

3.14 Report of the Exchange

The co-ordinator is responsible for the report of the exchange. The report of the age reading exchange might contain the following sections:

- abstract
- introduction
- material and methods
- results
- discussions
- conclusions
- recommendations.

Valid statistical tests and measures should be used to quantify the conclusions of the exchange. The co-ordinator should try to get firm conclusions concerning what preparation techniques or calcified structures to use (aim for standardising methods). He should discuss by mail the first draft of the report and incorporate the comments. Finally he should distribute the report to all participants and return the otoliths to the age reading laboratories.

3.15 Recommendation for an age reading workshop

A recommendation for an age reading workshop can only be given, if there is enough evidence of problems in age reading. These problems should be highlighted in such way that an age reading workshop can tackle these problems immediately.

3.16 After the exchange

If possible, when no serious differences in age readings were observed, digitised images should be produced of the calcified structures for which at least an 80% agreement was achieved. These can be used to start or to extend an "agreed collection", which is, and will become, a good tool for training new readers and for calibrating the age reading method of experienced and inexperienced age readers. See also sections 2 and 5.

4 AGE READING WORKSHOPS

4.1 Introduction

The main objective of an age reading workshop is to decrease the relative/absolute bias and to improve the precision (reduce CV) of age determinations (their reproducibility) between age readers of the different age reading laboratories.
An exchange of calcified structures should be carried out first to indicate the errors in age reading before a recommendation for an age reading workshop can be made (see previous section). The following possible problems in age reading might exist, e.g.

- the age reading methods of the age readers differ too much (as indicated by statistical tests);
- the precision in age reading is too low for certain age readers;
- there is a strong bias in the age readings of young and/or old fish;
- precision differs considerably for different preparation methods of the same calcified structures;
- inexperienced readers;
- other age reading problems.

At a workshop an attempt should be made to solve the problems indicated by the exchange.

4.2 Possible terms of reference for an age reading workshop

The following terms of reference could be formulated for the workshop:

- Review of the biology of the species (area of distribution of the population(s), time of spawning, time of feeding, maturity at age, maturity stages by season, growth in length or weight, migrations from spawning area to feeding area and to over-wintering area, etc.);
- Review of previous exchanges and workshops;
- Review when and how the age reading technique was validated;
- Review the sample processing techniques of the different age reading laboratories and try to standardise the processing techniques of calcified structures;
- Agreement on age determination criteria (e.g. date of birth 1st of January, one annual growth zone consists of one opaque and one translucent zone);
- Discuss disagreements in age reading results from the sets of the calcified structures read during the exchange and at the workshop and try to agree on the age reading method;
- Determine at the end of the workshop the precision in age reading and the relative bias (if possible the absolute bias);
- Estimate improvement in age reading concerning precision and bias by comparing exchange set and the last set at the workshop;

- Prepare a manual for age reading (date of birth, interpretation of rings and edges, period of opaque and translucent ring formation);
- Make recommendations on how to improve the age reading quality;
- Determine to what extend the age reading errors affect the assessment and environmental monitoring (recruitment, fishing mortality and spawning stock biomass);
- Indicate which calcified structures can be used for the "agreed collection" and if possible produce digitised images.

Other terms of reference might be formulated based on the conclusions from the exchange.

4.3 Age reading errors

The different age reading errors are described in 1.3.

4.4 Preparation of the workshop

Although the original recommendation for a workshop may have been derived from the results of an exchange scheme another recommendation might have to be obtained from an international organisation e.g. ICES, ICCAT, CECAF, EIFAC or EFAN etc in order to get funding for this workshop. For this purpose a chairperson, a venue and the meeting dates have to be agreed beforehand and a case formulated for the workshop.

The chairman of the age reading workshop should make a list of possible participants of which those who participated in the exchange should certainly be included.
The chairman should try to find a meeting place for the age reading workshop, where there are good facilities. The following should be taken into account:

- Enough good quality binocular microscopes to be available for all participants (if possible, the same type of microscopes commonly used by participants in their own age reading laboratory)
- Enough computers should be available, if age readings are carried out from digitised images;
- A high quality large screen for displaying the calcified structures during the plenary discussions;
- The meeting venue should be large enough for group discussions with facilities to view calcified structures, probably in front of a large screen. There must also be enough bench space for the actual age reading using microscopes and ancillary lighting. Space is also required for at least one computer and printer.
- An extra person to be available for entering age reading results in the spreadsheet, for photocopying, preparing statistical analysis etc.

The chairman of the workshop should inquire whether calcified structures of known age are available somewhere to be included as an extra set at the workshop (see section 3.4). If there are only a very limited number of validated calcified structures available, it can be included in an existing set and be indicated with a specific name in the column stratum of *Table 1* (the modal age for these calcified structures should

then be replaced by the validated age). In *Table 5* the results of this small sub-set of validated calcified structures will be presented (precision and absolute bias). In this small sub-set the comparison of the estimated mean age and the true age will show whether there is a serious problem in bias or not (see also section 3.4).
The requirements for each set of calcified structures are described in the section 3.3. The number of sets of calcified structures will depend on the number of age reading problems. In addition an extra set (or the exchange set) is needed at the end of the workshop to estimate the improvement in age reading concerning precision (CV) and relative or absolute bias by comparing the exchange set(s) and last set at the workshop. A warning should be given here that the chairman should make sure that he does not include too many sets, since there should be enough time available for discussing the age reading problems (most important issue at the workshop) and for preparing the manual on age reading.

Prior to the meeting the chairman could ask for help from an expert on microstructures, especially when age reading problems occur in the first couple of years of life, when false rings are likely to have been laid down. Technical advances in the field of microstructures enable experts to examine calcified structures and to calculate exactly when a ring is laid down in the first years of life. This work is invaluable in understanding ring formation and the attendance of a microstructure expert is to be strongly recommended. However, it should be taken into account that this expert can only be helpful if the work on microstructures has been carried out prior to the meeting.

Sets of calcified structures and input sheets for filling in the age reading results should be prepared prior to the meeting.

If necessary the chairman has to apply for funding for the workshop (see section 4.6.2).

4.5 Letter from chairman to workshop participants

The chairman of the workshop could put the following information in his letter to the participants:

- Mention the start and end dates of the workshop. Mention also the time of the start of the workshop meeting on the first day and the expected time of the end of the workshop (this is relevant for booking flights and hotel reservations). Try to arrange it in such way that participants are not arriving too late on the first day and are not leaving too early on the last day!
- The terms of reference (see section 4.2)
- Add, if needed, specific tasks to solve the existing problems in age reading e.g.
    - Ask a participant to provide calcified structures collected from the distinct modes in the length distributions of the catches by quarter or month to provide an indirect validation of the age readings, especially of the younger age groups;
    - Ask a microstructure expert to assist in solving the problems in the interpretation of the rings during the juvenile period. He should carry out this work prior to the workshop.

- Provide a description of the objectives of the workshop and provide an outline of the work schedule;
- Distribution of the tasks mentioned in the terms of reference in order to have at the start of the workshop the first drafts concerning sections on the review of the biology of the species, the descriptions of the processing methods, the manual on age reading, review of the validation, etc.;
- Try to allocate an expert in stock assessment to carry out the work on the estimation of the effect of age reading errors on the assessment as soon as possible immediately after the workshop meeting;
- Provide a first draft of the table of contents for the workshop report and indicate, per section, which persons are responsible for drafting the text for the sections;
- Distribute the report of the exchange in which the recommendation is made for a workshop and in which the age reading errors are described;
- Ask what kind of microscopes are used by the participants at their own age reading laboratories (if possible similar microscopes should be used at the workshop otherwise it is likely that there will be a reduction in the precision in age reading);
- Send a list of hotels with prices, addresses and phone/fax numbers and request participants to make their own hotel reservations (restrict the list to one or two recommended hotels in order to try to get as many participants as possible staying in the same hotel);
- If possible provide a map of the city on which the hotels and meeting place are indicated;
- Provide name, address and phone and fax numbers of the meeting place (if possible also the name of a contact person);
- Mention that participants should book their own flights in time for the meeting (see also section 4.6);
- Provide information on travelling facilities from airport to hotels and from hotels to the meeting place.

4.6 Travel arrangements

4.6.1 No funding of the workshop

The chairman has no specific responsibilities concerning travel arrangements when the travel costs and daily allowances are paid by the age reading laboratories to their own participants.

4.6.2 Funding of the workshop

The chairman is closely involved in the travel arrangements, if the workshop is funded, either wholly or partly, by an external agency e.g. EU. In such cases often 100% (or less) of the travel costs and daily allowances have to be paid to the participants. The chairman should encourage the participants to book cheap flights. At the end of the workshop the chairman has to ask the participants to complete a form in which the following has to be filled in:

a) time and date of both arrival and departure;

b) travel costs;

c) accommodation costs.

This form should be passed to the chairman together with all relevant receipts.

Receipts for hotel accommodation and meals are not necessary if the chairman has arranged in the contract for expenses to be paid at a flat rate *per diem*. After the meeting the chairman can pay the participants or the age reading laboratories the total amount for travel costs and daily allowance.

4.7 At the workshop

The workshop meeting is divided into the following stages:

4.7.1 Welcome

The chairman opens the workshop meeting and welcomes the participants. Eventually all the pre-planning results in the gathering of many like-minded individuals who are anxious to extend their expertise in age estimation. It should be remembered that many of the participants may not be accustomed to attending international meetings and initially they may feel insecure and lack self-confidence. The chairman should make every effort to re-assure them and explain that the success of the workshop is entirely dependent on their expertise. Check whether all readers understand and speak the language used at the workshop. If this is not the case possibly one of the other participants can help in translating as much as possible.

4.7.2 Working Schedule

The chairman should provide a working schedule for the whole workshop and outline the objectives of the workshop.

4.7.2.1 Presentation of biology and life cycle of species

The person(s), who prepared the first draft on the biology and life cycle, should give a short presentation.
Age readers are highly skilled in interpreting the image viewed under a microscope. However, this interpretation can be enhanced by knowing the life cycle and behaviour of the fish species, e.g. the area of distribution of the population(s), time of spawning, time of feeding, maturity, maturity stages by season, growth, migrations from spawning area to feeding area and to over-wintering areas, moving from pelagic stages to demersal stages, etc. Some of these events may help to understand why false or narrower rings are laid down. It is assumed that a review of this kind of information is useful for age readers. It should be indicated in what periods opaque material is laid down on the edges of the calcified structures taking into account that this is likely to be different by age group. It should also become clear to the age readers at what age thinner annuli are likely to occur due to spawning.

4.7.2.2 Presentation of the processing techniques by age reading laboratory

The co-ordinator or the participants should give a short presentation of the processing methods they use for their calcified structures. The co-ordinator should try to highlight the differences in the processing methods between age reading laboratories. Discussions should take place concerning the standardisation of the processing method. Possibly during the workshop different processing methods will be tested, but it would save a lot of time, if this could be undertaken prior to the workshop, as an exchange. If this was not done prior to the meeting, possibly a recommendation can be made at the workshop to exchange sets in order to examine which processing method achieves the best results in age reading.

4.7.2.3 Presentation by the expert on microstructures

Microstructure experts can help in the identification of annual rings in calcified structures, especially in younger ages of fish. An invited expert can present the results of his research on microstructures, which has been carried out prior to the workshop meeting. This will help in the discussions on the interpretation of the ring structures of the younger ages.

4.7.2.4 Definition of terminology

The chairman should give a clear explanation on the difference between precision and accuracy (see section 1.3) and explain other terms that will be used during the workshop (see section 7 for some suggested terms).

4.7.2.5. Presentation of results from previous exchanges and workshops

Valuable lessons can be learnt from previous exchanges and workshops. One of the participants should be tasked to review all previous activity on the species/stock involved and make a presentation at the workshop.

4.7.2.6 Presentation of the age reading problems of the exchange

An explanation should be given on the problems in age reading as observed during the exchange. For each age reader it should become very clear, what he is doing wrong and how he could improve his age reading method. Those calcified structures of the exchange set that caused problems in age reading can be recognised in *Table 1* as having a low percentage agreement and a high CV. Plenary discussions with the use of the large screen should take place to explain to each other where the annuli and the false rings lie. The age readers have to try to get a consensus. Once agreement is reached check whether this explanation agrees with the text of the first draft on the manual on age reading!

4.7.2.7 Manual on age reading

The author of the first draft on the manual on age reading should give a presentation after which an extensive group discussion should follow. During the course of the workshop this manual should be frequently improved e.g. after each plenary discussion on age reading comparisons. This manual can be used as a protocol for age reading this species as agreed by the age readers of the different age reading laboratories.

4.7.2.8 Age reading at the workshop

The chairman should ask the age readers to complete the input form and should mention that the essential information on month of catch, area of catch and length of the fish is provided on this input form (and sex if the growth is sex dependent). All age readers should get one input form for each set of calcified structures for entering their age reading results. The age readers have to enter one estimate of age per calcified structure, which then should be the best estimate. The programs for age reading comparisons can only handle one age estimate and not uncertain age estimates such as “6 or 7”. They should not use a + group (combining fish of the older ages in one group e.g. 10 years and older). The age readers should not omit any calcified structures for age reading, because it influences the results on precision and accuracy (the age readers are only comparable, if they all read both the easy and the difficult calcified structures). If the age readers have comments on certain calcified structures they should annotate the input form, referring to the identification number of the calcified structure. This is especially important when calcified structures become damaged at the workshop or when these are put in the wrong envelopes, etc. The chairman can exclude these from the final comparison of age reading results.
The chairman should request the participants to be extremely careful with the sets of calcified structures, since damage to the material will reduce the precision in age reading for subsequent age readers.

4.7.2.9 Age reading comparisons

A full explanation on age reading comparisons is given in section 3.13. Results of age reading comparisons (using the *ANALYSIS sheet of the Workbook*) should become available to the age readers within a few minutes after the last reader has finished the age readings of that set of calcified structures. This can be achieved by continuously inserting the age readings in *Table 1* as they become available during the age reading process. The most important component of the workshop comes once the data have been evaluated and disagreements located. Somebody has to present and evaluate the results to all age readers before a plenary discussion starts.
The most immediate statistical evaluations can be obtained from the spreadsheet and used in plenary discussion.. Further statistical evaluations (see section 6) can be undertaken at the workshop if resources permit or after the workshop has closed.

4.7.2.10 Plenary discussions on differences in age reading

The calcified structures, which caused large discrepancies in age reading, are easily found in *Table 1*. Large differences in age readings are indicated by a low percentage agreement and a high CV (low precision). A very helpful method is to display the calcified structures on a large screen on which the age readers can indicate how they counted the rings. Differences in age reading methods are likely to become evident. It is very useful to identify exactly where the differences in interpretation have occurred. The age readers have to try to get a consensus. Once agreement is reached check whether this explanation agrees with the text of the first draft on the manual of age reading!

4.7.2.11 Comparison of different preparation techniques or different calcified structures

See sections 3.8 and 3.9.

4.7.2.12 Age reading of the last set for estimating improvement in age reading

The last set of calcified structures read at the end of the workshop can be used to estimate the final age reading errors of the age readers (this could also be a set used during the exchange). The estimates of precision (CV), agreement and bias can be compared with those obtained from the exchange set. The CV's of both sets should be compared for each age group separately, if CV at age is not constant for all age groups and if the age compositions of both sets differ significantly. The results of both sets can be put in one table and it can be evaluated whether an improvement in age reading has been achieved at the workshop meeting (lower CV, higher agreement and lower bias).

Working conditions (e.g. unfamiliar microscopes), demanding time restrictions and adapting to new age reading methods can hamper the expected reader improvement.

Furthermore it is expected that the precision (CV) in age reading will not improve very much at the workshop, mainly because it is usual for only experienced age readers to participate in such workshops. These readers have a consistent age reading method (small spread in the age readings). This consistent age reading is only achieved after many years of experience and can not be learnt in a few days at a workshop. Probably only inexperienced readers could improve their precision at the workshop, and then only if they get appropriate guidelines from experienced readers on how to carry out the age reading.

At the workshop more improvement is expected regarding a decrease of bias and a increase in percentage agreement. Relative or absolute bias (accuracy error) provides information on the distance between mean age estimated and the modal/actual age. Readers, who have a bias in their age reading method, might be able to improve considerably in the percentage agreement to modal/actual age, if they learn at the workshop how to recognise the annual rings to be counted.

4.7.2.13 Preparation of an agreed collection

Based on the age readings, carried out on the last set for the determination of age reading errors at the end of the workshop, a reference set can be composed of calcified structures for which the age readers reached at least 80% agreement. This **agreed collection** can either be read in the next exchange in order to detect any divergence in the age reading method or can be used to start or to extend the **agreed collection**. The latter is an excellent tool for training new readers and for calibrating the age reading method of experienced and inexperienced age readers (see sections 2.3 and 5.7).

4.8 Effects of age reading errors on assessment and environmental monitoring

The effects of age reading errors on the assessment can be evaluated, if information on the precision at age (CV) and absolute bias at age are available. However, a similar exercise can be carried out when only information on relative bias at age is available under the assumption that relative bias is close to absolute bias (even a small collection of validated calcified structures could justify this assumption!).

The effects of age reading errors on the estimation of recruitment, fishing mortality and spawning stock biomass can then be studied. An example of such an evaluation is given in the addendum of the Report of the Horse Mackerel Otolith Workshop (ICES, 1999). However, such an evaluation will not be possible until after the workshop meeting, since the most up to date information on CV and bias at age will only become available at the end of the workshop meeting. This exercise ought to be inserted in the workshop report as an addendum.

4.9 After the workshop meeting

The chairman should remember that many of the participants may not be familiar with the requirements of an ICES workshop which dictate a comprehensive report. Thus much of the responsibility of writing and finalising the report will fall upon the chairman. The chairman should summarise the meeting and produce a report that reflects the work undertaken. It is therefore very helpful if any text that has been prepared prior to the meeting e.g. text on life history, can be agreed during the workshop to be included in the report. Valid statistical tests and measures should be used to quantify the conclusions of age reading workshops. It must be emphasised that the chairman will have to bear the bulk of this task, since the most valuable part of the workshop is the exchange of ideas and knowledge between recognised age readers. Too much text writing will dilute the time that they have available for the more important issue of age determination.

The chairman should distribute a preliminary report and should include the comments of the participants (unless text, tables and figures are agreed at the end of the workshop). Finally he should return the calcified structures and distribute the final report to the age reading laboratories.

All participants should be encouraged to remain in touch after the workshop and discuss on a one-to-one basis any future problems that may arise in their daily tasks.

5 DIGITAL IMAGING TOOLS

(Contributors: R. Ayers, M. Cunha, C. Nolan, A. Moreno, H Troadec)

5.1 Introduction

The use of digital images for age estimations of calcified structures can provide many benefits. Their ease of circulation and exchange should facilitate dialogue at national and international levels on methodology and interpretation protocols (see glossary of EFAN report 3/98).

The methods and technologies used in the capture and processing of digital images lie in a dynamic area that requires researchers to exchange information regularly. Hence it is supposed that the guidelines provided below will require frequent updating.

As a first step in the interpretation of calcified micro and macro structures, used in the determination of age, a detailed analysis of each structure is required. Analyses should identify and describe basic structures and their associated problems and incorporate these in the development of general, standardised interpretation procedures.

Resulting protocols should be objective and comparable, recognising intra and inter specific structural variability in their construction and incorporating recognised terminology in the formulation of instructions.

Such objectivity will assist in the resolution of subsequent, subjective interpretation of the image (e.g. subjectivity due to species life histories, geographical areas of occurrence) and the associated age estimation of the calcified structure.

5.2 Background

EFAN Cell 2 (Information Processing), a collaboration of software developers and users of digital image processing software, developed simple protocols for the exchange of interpreted digital images as a result of a workshop held in the University of Porto (EFAN, 1998). These protocols were used and adapted in a subsequent workshop on whiting otoliths where a reference collection of annotated images was created (Hirtshals (Denmark); ICES 1998). These protocols are in a continuous state of update and development and depend on continued communication between users and software developers for their successful application (EFAN, 2000).

5.3 Goals for digital image users

- The establishment of a digital image reference database of calcified structures (e.g. otoliths, statoliths, scales, spines etc.) of species aged with computer assisted age estimation tools (semi-automatic or automatic). Such a database should preferably be web based for maximum accessibility and include control, agreed and/or validated reference collections (e.g. *EFAN Cell 2 "Digital Image Reference DataBase of Interpreted Calcified Structures*"),

- The establishment of digital image exchange programmes using standardised protocols,
- The organisation of *ad hoc* workshops to resolve interpretation problems,
- The initiation, where possible, of dialogue between the users and developers of image analysis software.

5.4 Creation of a digitised reference collection

5.4.1 Software choice

The choice of a specific software package must be guided by the kind of use that is required (e.g. creation of the image set, interpretation of images, automatic post-processing of the results) and by the tools that are required to produce a correct interpretation (e.g. measurements, back-calculation, focusing). Software available on the market can be divided into three types:

1. Graphics packages or photo retouching packages providing tools enabling basic image and graphic overlay manipulation (e.g. Paintshop-Pro, PhotoShop). They allow the marking of basic structure information required for calcified structure interpretation. Digitisation is generally limited to TWAIN drivers that are not sufficiently flexible in most cases require the use of more specialist imaging packages.

2. Scientific imaging packages provide the features described in (1) along with sophisticated digitising capabilities and measurement tools. Some packages provide functions more or less adapted to age estimation on calcified structures (e.g. Image-Pro, NIH, Optimas, Visilog).

3. Specialised imaging software can provide, in addition to all the features described in (1) and (2), more specific functions like real-time back-calculation, ring location, curved profile acquisition, mosaic building (e.g. OTO, Bony-Parts, TNPC/Visilog). They are also able to manage basic structure co-ordinates, which can be of a great help for post-processing of large amounts of readings.

5.4.2 Image digitisation

The digital image of a calcified structure is only a representation of the 'operator's' preferred view for interpretation. The 'operator' may be a group of experts or a single technician. Associated with each reference collection, files should be provided detailing the source of the images and the method used to capture them. As a collection evolves, it is important that the images within that collection remain comparable regarding acquisition conditions.

5.4.3 Image resolution and colour/Black & White

Sensor resolution is constantly increasing but digital images are still not of a quality comparable to that observed by the human eye through a microscope. The specification of the otolith collection should attempt to provide images with the best resolution and the best dynamic range. Cameras providing images in excess of 1Mb are becoming more widely available and must be preferred to those of lower resolution. The choice of colour or black and white images should be made after discussion with expert readers who can give guidance on their use of colour gradients when making interpretations.

5.4.4 Viewing limitations

When an age reader is working with an original sample it is possible to alter all the viewing parameters e.g. lighting, orientation, magnification or focus; using digital images restricts the parameters that can be altered by each reader. This effectively restricts the reader to choices made by whoever digitised the sample.

Focusing problem
One of these limitations is that digitising only provides images in one focal plane. The choice of focal plane can vary according to the operator and can be considered as subjective. Today there is no solution available to this problem but ideally a sequence of images could be acquired while varying the focus i.e. in different focal planes. If measurements are required, a microscope with a motorised Z-axis will be necessary in order to precisely interpret this third dimension. This will produce a huge image file and its ease of manipulation will rely heavily on the memory (RAM) available on the computer. Additionally, it will require some programming and tests prior to validation and approval.

Limited field of view
An image may be a single captured image showing the whole structure or a mosaic of captured images (elements) joined to form the whole structure. When creating a mosaic image it is important that as many of the viewing parameters remain the same for each element of the mosaic. Mosaics can be built by manually overlaying each element (e.g. with a graphics or photo retouching package) or by using automatic mosaic builders, which attempt to find common areas. When using automatic mosaic building systems the stability of the viewing parameters is more important than that when building mosaics manually. Altering the viewing parameters can reduce the likelihood of an automatic system finding a correlation between elements.

Image format
Digitisation should be made in a format that does not result in degradation of image quality. A lossless image format is preferred e.g. TIFF LZW and users should be aware of the manner in which their software manages file compression.

Image ownership
With increasing use of digital images the ownership of an image may become an issue. For each collection of images, ownership should be detailed and their allowed use specified. Digital watermarking of images is a way of indelibly stamping ownership into an image without materially affecting it.

5.5 Image annotation

When collecting images for a reference collection the previous comments regarding operator preference should be given consideration (see 5.4.1 and 5.4.2).

The image collection will be composed of calibrated digital images with a resolution that still enables the reader to estimate the age. A graphic interpretation will be associated with each image in order to describe the location of each basic structure identified by the reader as being required for its age estimation. Rejected information produced by the interpretation process (e.g. false annuli) will be also included in the graphic information.

Large numbers of images can become a problem to manage. To aid the manager or co-ordinator an agreed file-naming convention should be adopted as one of the first steps in building an image collection. The convention should take into account the platforms across which the images may be used as some systems may not support extended filenames.

5.6 Image format

If no measurements or back-calculations are required from the image collection then Paintshop-Pro (version 5 and later) can be used as a tool for producing annotated images. Otherwise the selection of software will be governed by the tools required (see section 5.4.1), and similar procedures as described for Paintshop-Pro can be applied.

The PaintShop-Pro™ file format (version 5) (JASC Corporation, http://www.jasc.com) will be used as the file format with additional layers containing symbols and/or text annotation regarding the interpretation of the image

Graphic interpretation : Layer Descriptions

The layer identified as 'Background' contains the image of the calcified structure. When there is a need to merge layers from different files, a square reference point will be marked in an area away from the structure (Figure 5.6.1). This reference point must be duplicated in each new layer created by readers. The reference point allows accurate overlaying of marks from different readers into one image file for comparison.

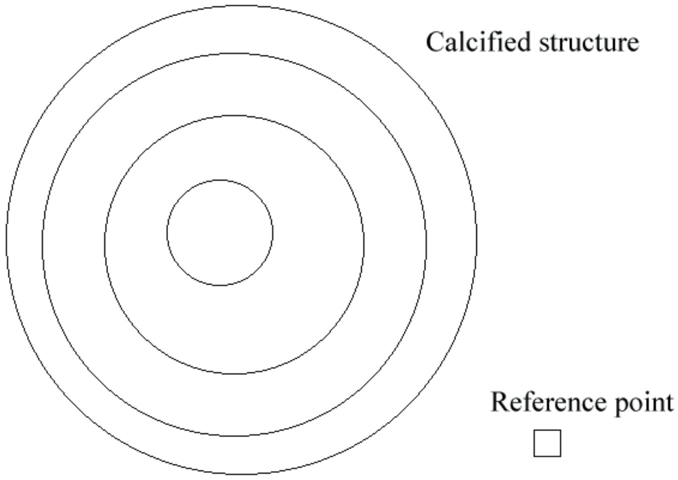


Figure 5.6.1: Reference point in the 'Background' layer.

Layer identified as "Nucleus", identifies the position of the nucleus with either a dot or a ring. (see Figure 5.6.3)

Layer identified as "Rings" contains the agreed checks or rings that the age reader should be attempting to identify. The rings can be identified with dots that contrast with the sample colours (see Figure 5.6.4)

Layer identified as "Info" highlights other features within the image that the reader feels would aid other readers when interpreting the visible structures e.g. false annuli. These identified features should be backed up with remarks in the text layer (see Figure 5.6.5)

Layer identified as "Text" contains remarks relating to features highlighted in the info layer and comments relating to the image and its interpretation (see Figure 5.6.6).

Co-ordinators should consider the facility to use different colour marks for the parameters they wish to study (Figure 5.6.2) e.g. For a between reader analysis a different colour would be used for each reader. Using different colours will also facilitate the post-processing of exchange exercises.

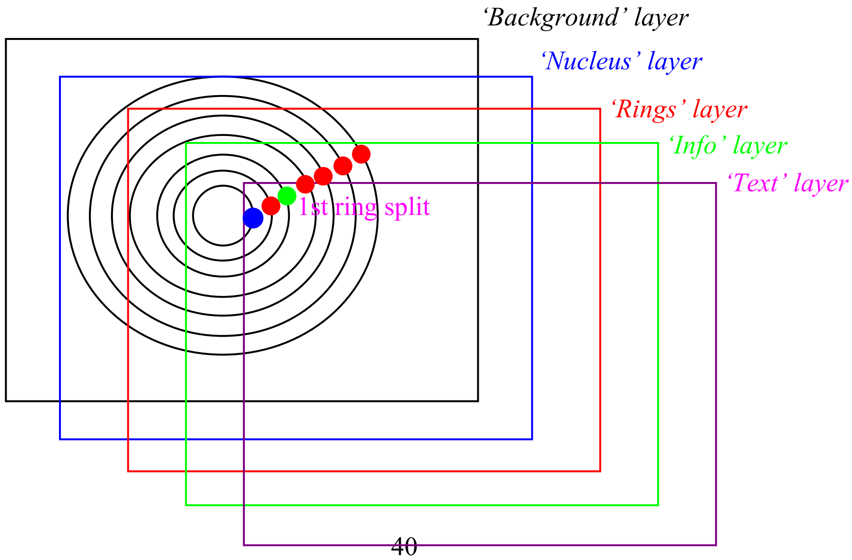


Figure 5.6.2: Schematic representation of a calcified structure image annotated with graphic layers.

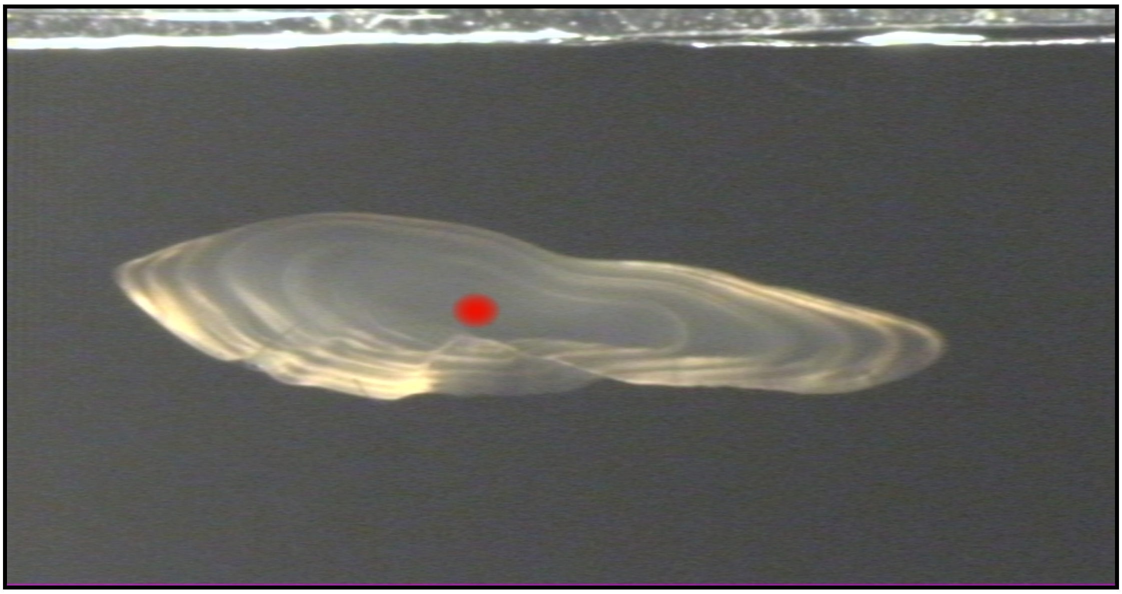


Figure 5.6.3 Image with nucleus marked (Large dot in centre)

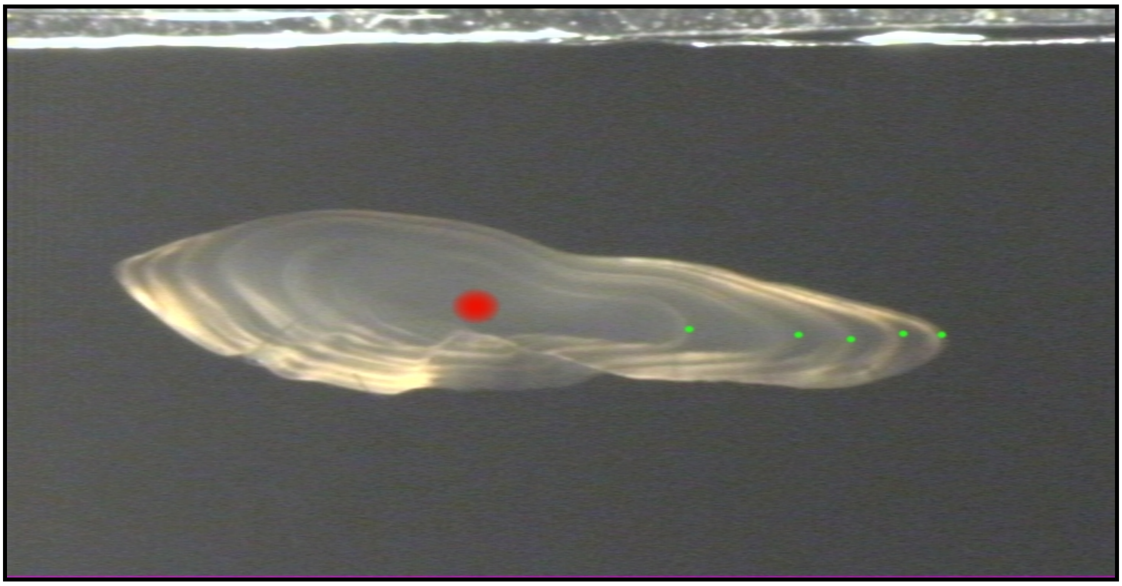


Figure 5.6.4 Image with nucleus (large dot) plus five annular rings (five small dots towards edge)

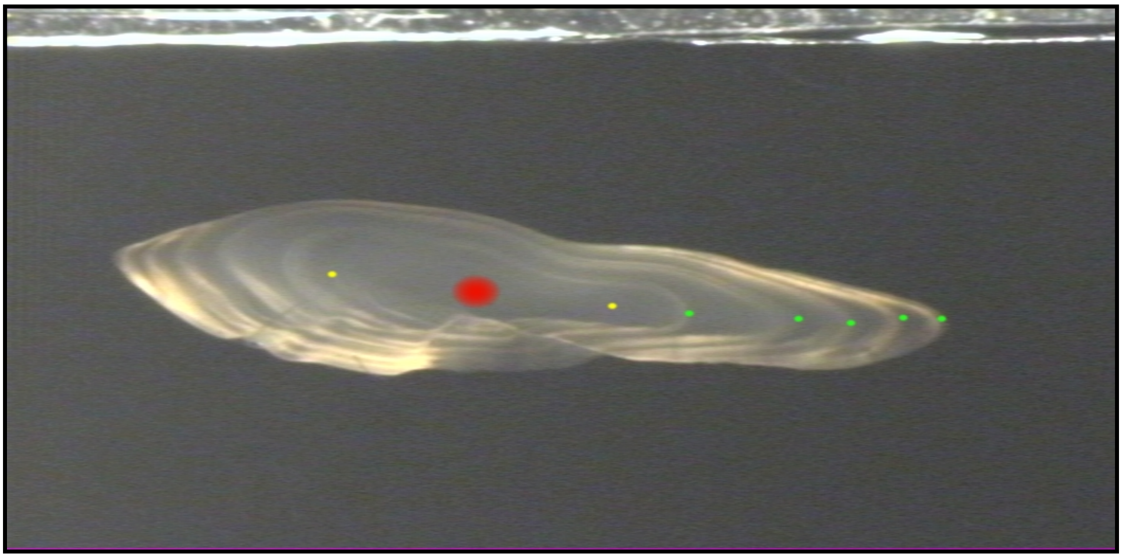


Figure 5.6.5 Image of 5.6.4 repeated with additional suspected false ring (1 small dot either side of nucleus)

Figure 5.6.6 Complete image with text explaining false ring and 5 annular zones

External data file
In addition to the images files, there will be a data file for each image set. This file contains information about the sample, how it was prepared and information on the equipment used to capture the image. The data file will be CSV format with one line per image file with 23 parameters, comma separated, on each line.

Data Required :

1. Submitting institute
2. Species: Latin name
3. NODC code
4. Institute assigned identifier -- Allows the original sample to be reprocessed if necessary.
5. Type of hard part (Otolith, Scale, Statolith, Spine, Vertebra, Tooth)
6. Length of individual (mm)
7. Sex of individual (M,F,U)
8. Maturity stage (ICES Scale plus U for Unknown)
9. Estimated age
10. Known age
11. Age units (Years, Days, Seasons)
12. Date of capture (DD/MM/YY)
13. Position of capture
14. Image presented by (Contact person in institute)
15. Preparation method of the sample
16. Lighting method (Reflected, Transmitted, Polarised, Filtered)
17. Camera type
18. Frame grabber type
19. Image resolution at capture
20. Magnification (pixel size or calibration for measurements)
21. Pre-processing completed before submission
22. Filename -- note possible problems with long filename support on some platforms
23. File size -- Kb

5.7 Use of digital images for control collections:

Digital control collections allow a greater flexibility for quality assurance (QA) procedures as they enable the control of a number of parameters that can affect an age estimate (e.g. lighting, sample degradation). The problem of sample degradation over time, which reduces precision, is solved with the use of a digital source. A digital file will remain static for as long as its storage media is viable. The image file would remain in a read-only state on a master device with copies being made each time the image is needed. By providing annotations in addition to age when interpreting a control collection the reader can give the QA manager a way to investigate the source of any detected imprecision.

The image format used for control collections will need, as a minimum, the 'Background' layer (i.e. the basic image of the structure) and the 'Rings' layers.

There must be also in place measures to keep track of information on the reader and the date of the estimation (e.g. this may be in the form of agreed filenames).

A larger control collection would reduce the problem of 'reader memory' retention (i.e. recognition of previous individual calcified structures and the recollection of prior age estimates) but requires greater manpower input at the beginning of the process to capture multiple images and provide a first estimate for each reader.

Although the viewing parameters were fixed when the image was captured it is possible for an age reader to adjust the image by the use of filters (e.g. brightness, contrast, etc). Due to the effect that digital filters can have on bias in the estimations, the use of such filters should be considered as part of the reading protocol.

When a reader is allocated a sample of digital images to read they will be given copies of the master image containing only the 'Background' layer. In the process of making their estimation, they will create a new 'Rings' layer duplicating the reference square. They will then mark their interpretation of the structure onto the new layer and save the file. Once they have completed all the samples required the marked files should be returned to the co-ordinator with their age estimations. If a problem in precision arises the co-ordinator can discuss the problem with the reader using their interpretation and historically marked images.

5.8 Use of agreed digital image collections for training

Because the images within an agreed collection contain the results of multiple interpretations of the structures they become useful to new readers to see a wider range of interpretations without the need for all the experienced readers to be present.

The image format for this type of collection should use all the layer types described above (i.e. section 5.6 *et seq.*). The 'Rings' layer should contain the agreed interpretation with additional 'Rings' layers containing the interpretation of each of the agreeing readers.

5.9 Validated digital image collections

Validated collections are an extremely valuable resource. Digital versions of validated collections would allow much wider use of these collections without the inherent risks involved in transporting original samples.

The image format for validated collections should include all the described layer types with text layer including information on or references to the method of validation.

5.10 Use of digital images for exchanges

One of the major problems with standard sample exchanges is the timelag as the samples are passed to each reader. In order to increase the speed of the exchange, database and software access might be proposed in different ways:

- By sending the database on CD-Rom (650Mb) or DVD-Rom (4.7 Gb),
- By downloading the database from an ftp site.

Co-ordinators should survey the participating institutes to identify a suitable level of hardware compatibility (e.g. the media type to be used for distribution, the operating platform to be used).

5.11 Use of digital images for workshops

Digitisation is the first stage in all uses of images and should form part of the introductory session of the workshop. The protocols used by individual institutes should be available for discussion and the co-ordinator should consider whether it is appropriate to digitise a sample of the structures during the first part of the workshop.

The images that will be used in the workshop should, where possible, be digitised to a previously agreed standard.

The co-ordinator should consider the use of networked facilities for holding the source images, in a read-only form, and the result images of the workshop sessions. These files must be included in routine secure backups.

At some point during the workshop it will become necessary to discuss images and results in a plenary session and facilities such as a large screen projector connected to a PC should be available.

If informations such as length, region of capture and season of capture are considered important factors in the estimation of age, it should be made available in an accessible form. It may be appropriate to include the information as part of each image.

5.12 Prospects for the future

5.12.1 **Adapted software requirement**

Using collection management software it would be possible to allocate samples to be read by a selected reader, pass those images to the reader, collect the resulting age estimations and compute agreement levels for constructing an agreed collection. The management software would assist the post-processing of the interpretation results (e.g. compare ring locations automatically, track and report changes in precision and provide additional analyses).

5.12.2 Intensification of internet use

A more intensive use of the internet would provide a solution for many current problems:

- A virtual workshop would ensure a quick and cheap collection of the data,
- Standard web browsers may be the solution to the hardware compatibility problems,
- Remote access to existing software via Web conferencing software (e.g. NetMeeting, Microsoft ™) or to specific Web-designed software would ensure a harmonisation of tools and protocols and eventually result in more efficient post-processing of the data.

6 FURTHER STATISTICAL PACKAGES

IPIMAR (the national fisheries institute of Portugal) is critically reviewing the statistical analysis procedures currently available and is also exploring the possibility of new statistical approaches.

In addition, IPIMAR, in collaboration with the Statistical Department of the Technical University of Lisbon, is trying to define a procedure for analysing results that will permit a more complete analysis of disagreement, particularly, in terms of evaluating trends. This work is ongoing and will take considerable time. In the interim period the authors endorse the recommendations from the 1994 ICES Report of the Workshop on Sampling Strategies for Age and Maturity (ICES, 1994). Co-ordinators should bear in mind that some of these procedures may require a considerable time to run and that the outputs may be confusing to some of the participants. It is recommended that an expert is available to perform and explain the statistical procedures as well as interpreting the results.

ICES (1994) presents a set of statistical tools that can be useful in the analysis of data from age reading workshops or exchanges. This given set is by no means exhaustive or optimal, but should provide a reasonable basis until refinements are devised. A short summary of the statistical tools as recommended in ICES (1994) is listed below:

- <u>Testing for between-reader bias</u>: The t-test and the Wilcoxon rank sum test always give the same classifications. Bowker's test (Bowker, 1948) is usually less powerful, however, except when asymmetry doesn't cause bias, for example when a small number of large negative errors is compensated by a large number of small positive errors or when negative and positive errors are associated with different ages.

- <u>Measurement agreement between unbiased age readers</u>: Three measures of agreement can be considered: the average percent age error (APE) (Beamish and Fournier, 1981), the coefficient of variation (CV) (Chang, 1982) and the chance-corrected observer-agreement measure (kappa) (O'Connel and Dobson, 1984; Schouten, 1982). The first two measures are relative estimates of variability, the third is a true measure of inter-reader agreement, corrected for the level of agreement that would be expected even if the readers were assigning ages at random (for the same common age distributions).

- <u>Testing for group membership</u>: Each member of an unbiased group of age readers can be tested for membership in that group. This can be detected by calculating the likelihood of a reader's choices among all the observed readings. This likelihood, which is normally distributed under the hypothesis that the members of the group are interchangeable, can be compared with its expected value. Likelihood values that are too small indicate members of a group that do not agree often enough with the rest, while likelihood values that are too large indicate membership of a tight subgroup.

* Estimating dispersion parameters: The probability models and the maximum likelihood methods described in Richards *et al*. (1992) can be used to estimate the reading dispersion rates of the readers in a homogeneous group. If the dispersion rates of readers are known, it is then possible to correct for the effect of ageing dispersion on age distribution estimates (Richards *et al*. (1992)).

7 GLOSSARY FOR CALCIFIED STRUCTURE STUDIES

The terminology used in this report is taken from the glossary of the first "Fish Otolith Research and Application" symposium presented in Secor *et al*. Terms added by the authors are highlighted in bold.

Absolute bias – the systematical over- or underestimation of age compared to true age.

Accessory growth centre – a growth centre formed beyond the otolith core that leads to a new plane of growth and from which a new series of growth increments appears to emanate. Formation of these structures is often associated with life-history transitions such as metamorphosis. Accessory growth centres are often referred to as accessory primordia; however, the term accessory growth centre is preferred because these features are different structurally from primordia (e.g., they do not contain primordial granules).

Accuracy – the closeness of a measured or computed value to its true value (how close the estimated ages are to the true ages). Bias is the over- or underestimation of age. Absolute bias is in comparison to true or known age and relative bias is in comparison to modal age.

Age estimation, age determination – these terms are preferred when discussing the process of assigning ages to fish. The term ageing should not be used as it refers to time-related processes and the alternation of an organism's composition, structure, and function over time. The term age estimation is preferred.

Age group – the cohort of fish that has a given age (e.g. the 5-year-old age group). The term is not synonymous with year-class or day-class. All ages are assumed to begin on January 1st.

Agreement – the percentage of age readings that agree with other age readings or with the modal or true ages. The same age reader can estimate the agreement from repeated age readings. The agreement depends on age and should preferably be expressed as agreement at age.

Annulus (pl. annuli) – one of a series of concentric zones on a structure that may be interpreted in terms of age. The annulus is defined as either a continuous translucent or opaque zone that can be seen along the entire structure or as a ridge or a groove in or on the structure. In some cases, an annulus may not be continuous or obviously concentric. The optical appearance of these marks depends on the otolith structure and the species and should be defined in terms of specific characteristics on the structure. This term has traditionally been used to designate year marks even though the term is derived from Latin "anus" meaning ring; not from "annus" which means year. The variations in microstructure that make an annulus a distinctive region of an otolith are not well understood.

Average age – the average age computed by calcified structure, based on the ages attributed by all age readers.

Between-reader-bias – the difference between the age readings of two age readers

Bias – the systematical over or underestimation of age. Bias is regarded as a serious error since fish are allocated to the wrong year class.

Check – a discontinuity (e.g., a stress-induced mark) in a zone, or in a pattern of opaque and translucent zones, or microincrements. Microstructural checks (e.g., hatching checks) often appear as high-contrast microincrements with a deeply etched D-zone or an abrupt change in the microstructural growth pattern. If the term is used, it requires precise definition.

Cohort – group of fish of a similar age that were spawned during the same time interval. Used with age group, year-class, and day-class.

Core - the area or areas surrounding one or more primordia and bounded by the first prominent D-zone. Some fishes (e.g., salmonids) possess multiple primordia and multiple cores.

Corroboration – a measure of the consistency or repeatability of an age determination method. For example, if two different readers agree on the number of zones present in a hard part, or if two different age estimation structures are interpreted as having the same number of zones, corroboration (but not validation) has been accomplished. The term verification has been used in a similar sense; however, the term corroboration is preferred as verification implies that the age estimates were confirmed as true.

D-zone – that portion of a microincrement that appears dark when viewed with transmitted light, and appears as a depressed region when acid-etched and viewed with a scanning electron microscope. This component of a microincrement contains a greater amount of organic matrix and a lesser amount of calcium carbonate than the L-zone. Referred to as a discontinuous zone in earlier works on daily increments; D-zone is the preferred term. See L-zone.

Daily increment – an increment formed over a 24-hour period. In its general form, a daily increment consists of a D-zone and an L-zone. The term is synonymous with "daily growth increment" and "daily ring". The term daily ring is misleading and inaccurate and should not be used. The term daily increment is preferred. See increment.

Discontinued ring – A ring read as an annulus but which cannot be followed over the entire calcified structure. These rings often appear to be fused to the preceding annulus and in whiting otoliths, for example, are restricted to fish > 3 years old.

Distal surface – opposite surface of the otolith from the *sulcus acusticus*.

False ring – a ring, which should not be interpreted as an annual ring, because it is laid down on the calcified structure as a result of random, short-term fluctuation in some environmental parameter briefly affecting growth rate rather than a regular seasonal change in the environment.

Growth pattern – in several species size of annuli vary according to a pattern. This pattern when considered as a stable one can be used to distinguish between true and false annuli for the first year's growth.

Hyaline zone – a zone that allows the passage of greater quantities of light than an opaque zone. The term hyaline zone should be avoided; the preferred term is translucent zone.

Increment – a reference to the region between similar zones on a structure used for age estimation. The term refers to a structure, but it may be qualified to refer to portions of the otolith formed over a specified time interval (e.g. subdaily, daily, annual). Depending on the portion of the otolith considered, the dimensions, chemistry, and period of formation can vary widely. A daily increment consists of a D-zone and an L-zone, whereas an annual increment comprises an opaque zone and a translucent zone. Both daily and annual increments can be complex structures, comprising multiple D-zones and L-zones or opaque and translucent zones, respectively.

L-zone – that portion of a microincrement that appears light when viewed with transmitted light, and appears as an elevated region when acid etched and viewed with a scanning electron microscope. The component of a microincrement that contains a lesser amount of organic matrix and a greater amount of calcium carbonate than the D-zone. Referred to as incremental zone in earlier works on daily increments: L-zone is the preferred term. See D-zone.

Lapillus (pl. lapilli) – one of the three otolith pairs found in the membranous labyrinth of osteichthyan fishes. The most dorsal of the otoliths, it lies within the *utriculus* ("little pouch") of the pars superior. In most fishes, this otolith is shaped like an oblate sphere and it is smaller than the *sagitta.*

Marginal increment – the region beyond the last identifiable mark at the margin of a structure used for age estimation. Quantitatively, this increment is usually expressed in relative terms, that is, as a fraction or proportion of the last complete annual or daily increment.

Microincrement – increments that are typically less than 50 µm in width; the prefix "micro" serves to indicate that the object denoted is of relatively small size and that it may be observed only with a microscope. Often used to describe daily and subdaily increments. See increment.

Microstructural growth interruption – a discontinuity in crystallite growth marked by the deposition of an organic zone. It may be localised or a complete concentric feature. See check.

Modal age – the age for which most readers have a preference. It is recommended that the modal age is based on the age readings of the experienced readers.

Nucleus – in this report the term is used to denote the early part of the sagitta otolith formed prior to any accessory growth centres. This part of the otolith approximately corresponds to the otolith formed during the larval phase.

Opaque zone – a zone that restricts the passage of light when compared with a translucent zone. The term is a relative one because a zone is determined to be opaque on the basis of the appearance of adjacent zones in the otolith (see translucent zone). In untreated otoliths under transmitted light, the opaque zone appears dark and the translucent zone appears bright. Under reflected light the opaque zone appears bright and the translucent zone appears dark. An absolute value for the optical density of such a zone is not implied. See translucent zone.

Precision – the variability in the age readings. In the age bias plots the precision is preferably indicated by the mean age ±2stdev, while in general as the coefficient of variation (CV) by age group or as a weighted mean over all age groups.

Primordium (pl. primordia) – the initial complex structure of an otolith, it consists of granular or fibrillar material surrounding one or more optically dense nuclei from 0.5 µm to1.0 µm in diameter. In the early stages of otolith growth, if several primordia are present, they generally fuse to form the otolith core.

Proximal surface – *sagitta* surface facing the sensory macula, distinguished by having a *sulcus acusticus*.

Reader bias – the subjective assignment of ages, which is caused by the knowledge of existing strong or weak year classes. This causes strong year classes to become stronger and weak year classes to become weaker.

Relative bias – a systematical over- or underestimation of age compared to the modal age.

Sagitta (pl. *sagittae*) – one of the three otolith pairs found in the membranous labyrinth of osteichthyan fishes. It lies within the *sacculus* ("little sack") of the pars inferior. It is usually compressed laterally and is elliptical in shape; however, the shape of the sagitta varies considerably among species. In non-ostariophysan fishes, the *sagitta* is much larger than the *asteriscus* and *lapillus*. The *sagitta* is the otolith used most frequently in otolith studies.

Split rings - in many species the annuli have a tendency to split and appear as several annuli very close together. This causes difficulty in determining which rings are annuli and which are false rings. Ring splitting may be due to multiple spawning events during the year especially in species like whiting which are serial spawners. Reader experience and knowledge of the growth pattern is used to overcome the difficulties of grouping split rings.

Sulcus acusticus (commonly shortened to sulcus) – a groove along the medial surface of the sagitta. A thickened portion of the otolithic membrane lies within the sulcus acusticus. The sulcus acusticus is frequently referred to in otolith studies because of the clarity of increments near the sulcus in transverse sections of sagittae.

Transition zone – a region of change in structure between two similar or dissimilar regions. In some cases, a transition zone is recognised due to its lack of structure or increments, or it may be recognised as a region of abrupt change in the form (e.g.

width or contrast) of the increments. Transition zones are often formed in otoliths during metamorphosis from larval to juvenile stages or during significant habitat changes such as the movement from a pelagic to a demersal habitat or a marine to freshwater habitat. If the term is used, it requires precise definition.

Translucent zone – a zone that allows the passage of greater quantities of light than an opaque zone. The term is a relative one because a zone is determined to be translucent on the basis of the appearance of adjacent zones in the otolith (see opaque zone). An absolute value for the optical density of such a zone is not implied. In untreated otoliths under transmitted light, the translucent zone appears bright and the opaque zone appears dark. Under reflected light the translucent zone appears dark and the opaque zone appears bright. The term hyaline has been used, but translucent is the preferred term.

Transverse section – section through the proximal-distal axis.

Validation – the process of estimating the accuracy of an age estimation method. The concept of validation is one of degree and should not be considered in absolute terms. If the method involves counting zones, then part of the validation process involves confirming the temporal meaning of the zones being counted. Validation of an age estimation procedure indicates that the method is sound and based on fact.

Vateritic – term used to describe otoliths in which all or part of the calcium carbonate has been deposited in the vaterite rather than the more typical aragonite form. Vateritic sagittae appear ‘glassy’, making it difficult to differentiate incremental structures. The cause of vateritic replacement is unknown but has been found to be common in fish kept under hatchery, and therefore perhaps stressed, conditions.

Verification – the process of establishing that something is true. Individual age estimates can be verified if a validated age estimation method has been employed. Verification implies the testing of something, such as a hypothesis, that can be determined in absolute terms to be either true or false.

Year-class – the cohort of fish that were spawned or hatched in a given year (e.g., the 1990 year-class). Whether this term is used to refer to the date of spawning or hatching must be specified as some high-latitude fish species have long developmental time prior to hatching.

Year-class-strength bias - the effect of precision errors in age reading on the estimated age composition due to differences in year class strength, which results in an overestimation of weaker year classes and an underestimation of the stronger ones (smoothing effect).

Zone – region of similar structure or optical density. Synonymous with ring, band and mark. The term zone is preferred.

8 Acknowledgements

Whilst the authors have drawn on their own experiences they are also indebted to the many other fishery scientists who have contributed to the establishment of these guidelines, in particular to members of EFAN Cell 2 who provided extensive guidelines on digital images (section 5). We are also indebted to those members of Cells 2, 3 and 4 of EFAN who spent many hours reviewing this manuscript at the 5th EFAN plenary session in Palma, Mallorca during October 2000. Finally thanks are due to Dr Conor Nolan of Trinity University, Dublin for an editorial review of the document and for his suggested improvements in the original style and grammar.

9 References

Beamish, R.J. and Fournier, D.A. 1981. A method for comparing the precision of a set of age determinations. Can. J. Fish. Aquat. Sci. 38: 982-983.

Bowker, A.H. 1948. A test for symmetry in contingency tables. J. Amer. Stat. Assoc. 43: 572-574

Campana, S.E., Annand, M. C. and McMillan, J.I. 1995. Graphical and Statistical Methods for Determining the Consistency of Age Determinations. Transactions of the American Fisheries Society: 124: 131-138, 1995.

Chang, W.Y.B. 1982. A statistical method for evaluating the reproducibility of age determination . Can. J. Fish. Aquat. Sci. 39: 1208-1210

EFAN, 1998. Report of the EFAN Cell 2 Workshop on the " Establishment of a digital image reference database" held at the University of Porto, 28-29 September 1998, EFAN Report 3/98, Editor: H. Troadec, 42 p.

EFAN, 1999. Report of the EFAN Cell 2 Workshop on the Contrast enhancement and pre-processing of calcified structure images" held at the Candia Maris Hotel, Heraklion, Greece, 4-5 October 1999, EFAN Report 1/2000, Editor: H. Troadec, available at: http://www.efan.no/rep2000/C2Rep99.htm.

Francis, R.I.C.C. 1995. The analysis of otolith data – A mathematician's perspective. *In* Recent developments in fish otolith research. Secor, D.H., Dean J.M. and Campana, S.E. (Eds.). University of South Carolina Press. Columbia, South Carolina, USA, 81-96, 1995.

ICES 1994. Report of the Workshop on Sampling Strategies for Age and Maturity. ICES CM 1994/D:1, 67 pp.

ICES 1995. Report of the Workshop on Mackerel Otolith Reading. ICES CM 1995/H:1, 45 pp.

ICES 1998. Report of the Workshop on Otolith Ageing of North Sea Whiting. ICES CM 1998/G:14, 66 pp.

ICES 1999. Report of the Horse Mackerel Otolith Workshop. ICES CM 1999/G:16, 80pp.

Liew, P.K.L. 1974. Proc. int. Symp. on Ageing of Fish. Ed T.B Bagenal. Unwin Brothers.124-136, 1974.

Lin.L.I-K., 1989. A Concordance Correlation Coefficient to Evaluate Reproducibility. Biometrics 45,225-268

O'Connell, D.L., and Dobson, A.J. 1984. General observer-agreement measures on individual subjects and group of subjects. Biometrics 40: 973-983

Richards, L.J., Schnute, J.T., Kronlund, A.R., and Beamish, R.J. 1992. Statistical models for the analysis of ageing error. Can. J. Fish. Aquat. Sci.48: 1801-1815.

Schouten, H.J.A., 1982. Measuring pairwise interobserver agreement when all subjects are judged by the same observers. Statistica Neelandica 36: 45-61

Secor, D.H., Dean J.M. and Campana, S.E. (Eds.) Recent developments in fish otolith research. University of South Carolina Press. Columbia, South Carolina, USA, 723-729, 1995

Appendix 1

First sheet of 'Age Comparisons.xls' (READ ME FIRST)

A.T.G.W. Eltink

Version: 1.0 October 2000

Guus Eltink, RIVO, P.O.Box 68, 1970 AB IJmuiden, the Netherlands
e-mail address: a.t.g.w.eltink@rivo.wag-ur.nl
Direct phone: +31 255 564691

This Excel workbook ("AGE COMPARISONS.XLS") is developed for an easy and fast analysis of age reading results, e.g. at age reading workshops the results on the age reading comparisons can be provided immediately after the age readers have finished age reading. The results from the analysis on age reading comparisons are easy to understand for the age readers, who have to calibrate their age reading method based on these results. Furthermore this tool for age reading analysis is easy to adapt e.g. to analyse different types of reference collections, to compare to known age in stead of modal age, etc.

The essential companion to this Excel workbook is the document GUIDELINES AND TOOLS FOR AGE READING COMPARISONS by A.T.G.W. Eltink, A.W. Newton, C. Morgado, M.T.G. Santamaria and J. Modin, because it refers to this workbook for the analysis of age reading results. This document is an attempt to formalise and advertise the best features on age reading comparisons currently available in Europe. The guidelines concentrate on:

* Explanation of age reading errors;
* Reference collections and the analysis of the age determinations;
* Running exchange schemes and the analysis of the age determinations;
* Organising workshops and the analysis of the age determinations;
* Definition of terms.

Especially section 3.13 (Age reading comparisons) is relevant for the explanation and interpretation of the tables and the figures of this Excel workbook.

Both the file "AGE COMPARISONS.xls" and file "GUIDELINES.pdf" can be downloaded as one zip-file called AGE COMPARISONS.ZIP from the internet http://www.efan.no or can be obtained from the author.

You are kindly requested to refer to the workbook AGE READING COMPARISONS as follows:
Eltink, A.T.G.W. 2000. Age reading comparisons. (MS Excel workbook version 1.0 October 2000) Internet: http://www.efan.no

And to the document GUIDELINES AND TOOLS FOR AGE READING as follows: *Eltink, A.T.G.W., A.W. Newton, C. Morgado, M.T.G. Santamaria and J. Modin, 2000. Guidelines and tools for age Reading. (PDF document version 1.0 October 2000) Internet: http://www.efan.no*

It would be very much appreciated, if you would provide comments in order to further help improve the guidelines and the workbook for the analysis of age readings. Please download the latest version each time when you have to analyse the age readings of exchanges, of workshops or of control/agreed/validated collections.

Possible changes to sheet "ANALYSIS"

The number of 200 calcified structures in the set can be reduced by deleting the empty rows. The number of calcified structures in the set can be increased by inserting the needed number of rows and by dragging down the formulas of columns W, X and Y of Table 1. Do not forget to drag down the formulas of columns AB, AD and AL (for mean age, 80% agreement and the bias work table). Put "-" in the cells of Table 1, where there are no age readings or fish sample data.
When the age reading results of, for instance 9 age readers, have to be compared, the columns for age readers 10 - 15 should be deleted. Furthermore you have to delete the bias plot figures for those readers that are removed.
When the age reading results of 25 age readers have to be compared, then 10 columns have to be inserted (before age reader 15 and not after otherwise these data will not be included in the calculation of modal age, agreement and CV by calcified structure!). In all tables below Table 1 the formulas have to be dragged to the right. Additional bias plot figures have to be made.
The number of modal age groups in the age reading comparisons can be reduced by deleting in the tables the rows for those modal age groups that are not needed e.g. 13-15. The number of modal age groups in the age reading comparisons can be increased by inserting in all tables the number of rows for those age groups that are needed and by dragging down the formulas. The table on "Inter-reader bias tests" (Table 3) has to be filled in manually! The figures of the age bias plots should be changed in such way that the maximum value of y-axis (estimated mean age) is set to a value somewhat higher than the maximum value of the x-axis, because the highest mean age +2stdev must remain within the range of the y-axis.

Instructions concerning input data, calculation, printing etc. of the "ANALYSIS" sheet

Fill in the blue cells of the spreadsheet. First insert a table number in cell A1 and a title into cell D1, and the fish sample information into cells A5:G204. Columns A and G of Table 1 should be used only for stratum and month and should not be used for anything else, because an analysis by month and stratum is carried out (see Table 4 and 5). Fill in the reader names and countries in H3:V3. Insert the results of the age readings in cells H5:V204. If you have no age readings please insert a "-". Warning: do not use "cut" and "paste", because it will change the formulas! The names of the strata given in blue in Table 5 (cells AI699:AT699) should exactly correspond to the names of the strata as used in column A of Table 1.

The modal age is calculated automatically from the age readings of each calcified structure. However, sometimes two modal ages are possible, if a certain number of readers prefers one age and a same number of readers prefers another age. The excel speadsheet will use the first calculated modal age as the final modal age. Therefore, the most experienced readers should be put in the left hand columns (with lowest CV and highest agreement) and the most inexperienced readers on the right hand side (with highest CV and lowest agreement). In this way the modal age will be based on the most experienced readers, if two modal ages are possible.
If a number of inexperienced age readers participate, it is recommended to make the modal age to be determined by a selected group of experienced readers and not by the whole group!!! You should mention in your report on age reading comparisons, over which age readers you calculated the modal age and on which criteria you based this decision. In Table 1 you have to change in the heading the range of readers that you use for the calculation of modal age and you have to change the formulas for the calculation of modal accordingly for this range of readers!
Furthermore there can be a problem with the estimation of the modal age, if e.g. all age readings differ and no modal age can be estimated. This is likely to happen when the number of age readers is low (3 or 4 readers). This problem might be solved by inserting the rounded mean age in the cell that causes the problem (for that reason column "Mean age" is added to the right hand side of Table 1).

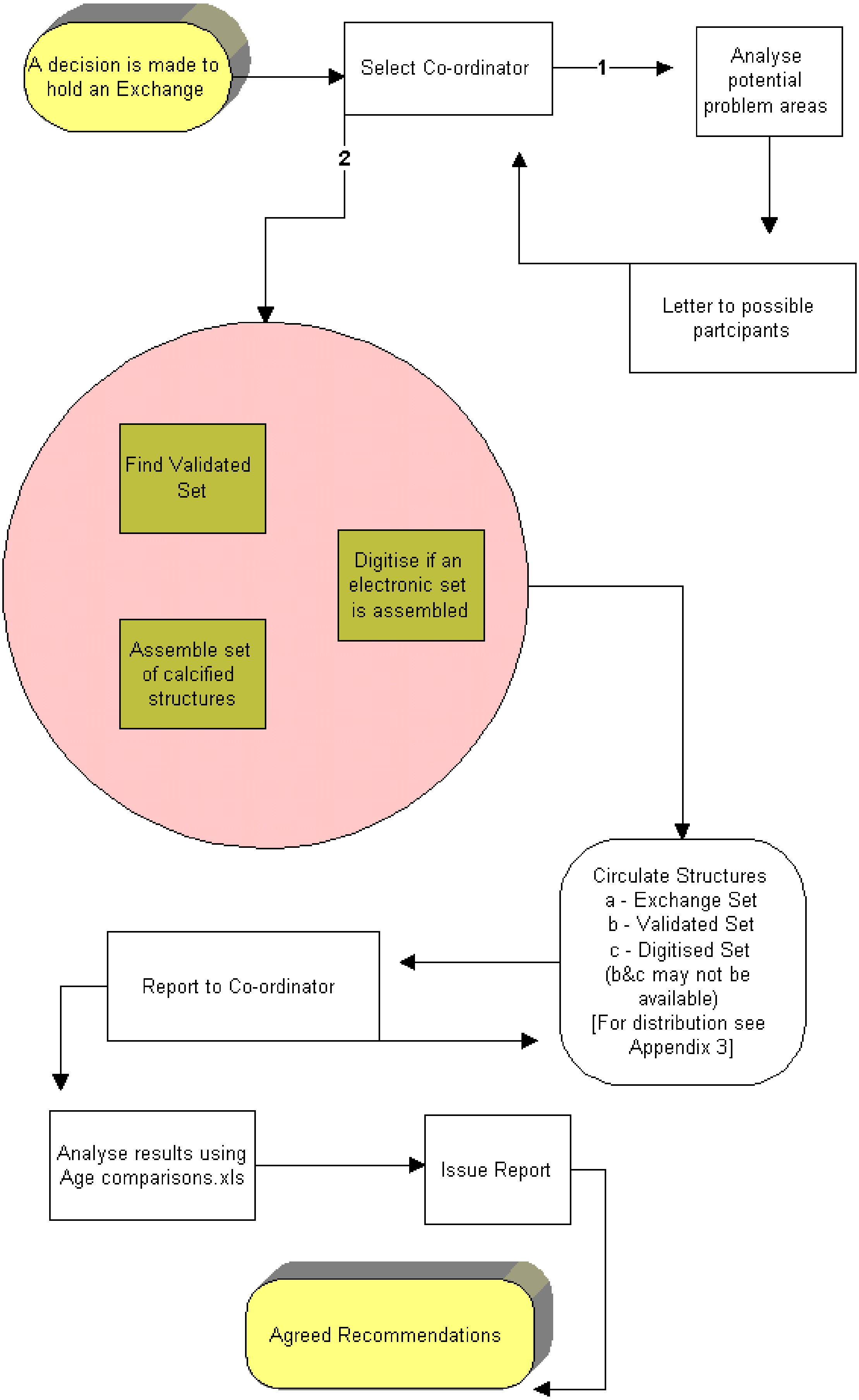
Execute "print" will provide you with output of the following tables and figures:

- Table 1: Basic input data with modal age, agreement and CV by calcified structure.
- Table 2: By age reader and by modal age the number of age readings, CV(%), agreement (%), relative bias and ranking of age readers.
- Table 3: Age composition and mean length at age by reader; inter-reader and reader against modal age bias tests.
- Table 4: By month and by modal age the number of age readings, percentage agreements, CV's and relative bias.
- Table 5: By stratum and by modal age the number of age readings, percentage agreements, CV's and relative bias (stratum is e.g. otoliths from different institutes or from different processing methods or a combination of certain months etc.);
- Table 6: Tables for plotting the age bias plots: 2stdev, mean age, 2stdev+mean age and 2stdev-mean age by modal age;
- Table 7: Overview of the number of calcified structures by modal age for which at least 80% agreement was reached.

- Figure 1: Age bias plot figures.
- Figure 2: CV(%), agreement(%) and standard deviation plotted against modal age.
- Figure 3: The distribution of the age reading errors in percentage by modal age.
- Figure 4: The relative bias by modal age.
- Figure 5: The estimated mean length at age by age reader.

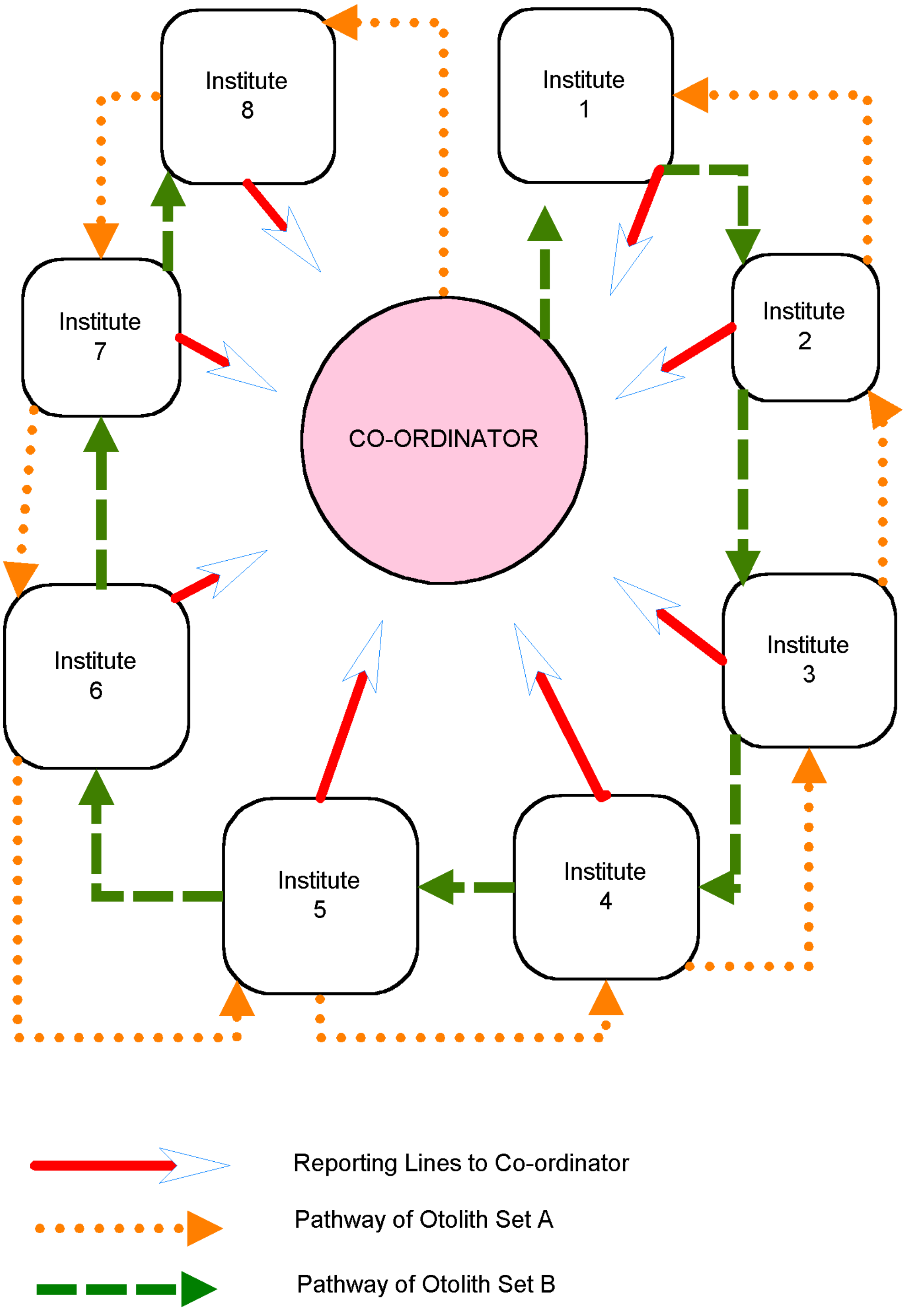
The quality of the black and white prints of the tables and figures will be improved, if the colours are changed to black.

Appendix 2 Exchange Flowchart



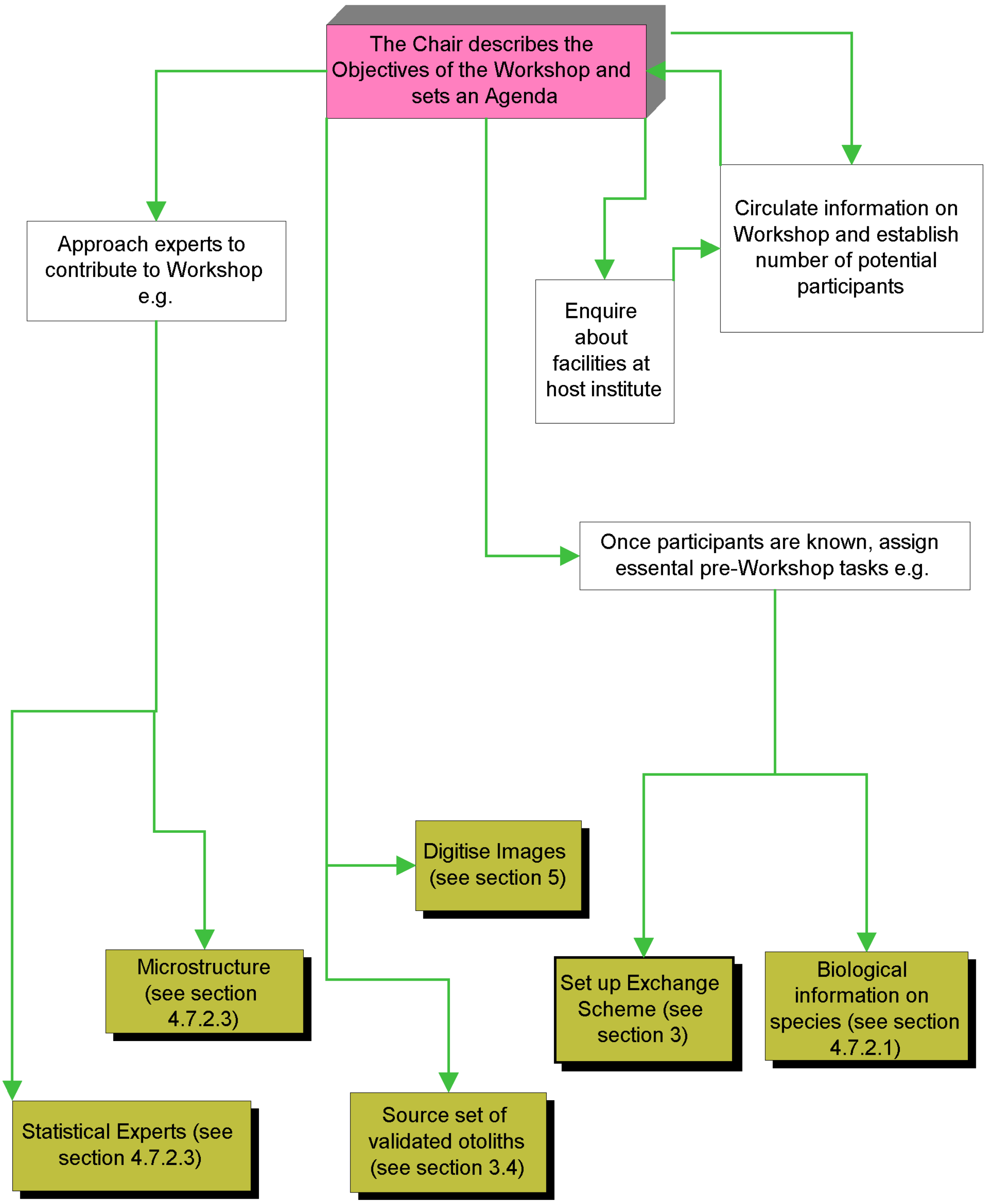
Appendix 3 - Distribution Chart

This Flowchart demonstrates a two-way distribution that could be used if the Exchange scheme had a large number of participants. The otoliths to be read are divided into two sub-sets - A and B.



Appendix 4a - Flowchart for Planning an Otolith Reading Workshop

The following items should be started at least 12 months before the Workshop



Appendix 4b - The Workshop