Northwest Atlantic   Fisheries Organization

SCIENTIFIC COUNCIL MEETING – JUNE 2002

Results of the Greenland Halibut (*Reinhardtius hippoglossoides*) Otolith Exchange
Between Spain, Canada and Portugal.

by

R. Alpoim[1], E. Roman [2], B. Greene[3], R. Burry[3] and W.R. Bowering[3]

[1] Instituto de Investigação das Pescas e do Mar, Av. Brasília 1400 Lisboa, Portugal.
[2] Instituto Español de Oceanografía, P.O. Box 1552, Vigo, Spain.
[3] Dept. of Fisheries and Oceans, Science, Oceans & Environment Branch,
P.O. Box 5667 St. John's, NF, Canada

Abstract

This Greenland halibut (*Reinhardtius hippoglossoides*) otolith exchange was carried out in order to examine for differences in age interpretations between Spain, Canada and Portugal. Statistical and graphical methods were used to evaluate differences in terms of bias and precision. In general, most differences in interpretations among age readers were within +/- 1-2 years although there were strong biases in several cases. The Spanish age reader had a tendency to assign lower ages to fish younger than 9 years old and higher ages to the older ones, relative to the Canadian age readers. The age reader from Portugal, like the Spanish age reader, assigned higher ages to the older fish (>8 years old) than Canada, but for the younger fish the agreement was very good. Burning the otoliths before reading seems to enhance the clarity in the last few annuli in older fish.

Introduction

In recent years, the assessment of the Subarea 2 and Divisions 3KLMNO Greenland halibut (*Reinhardtius hippoglossoides*) stock, age length keys from the Canadian fleet were applied to the length frequency of all fleets, instead of using the corresponding fleet age length key. This was due to the lack of agreement among the Greenland halibut age readers from several countries that have fleets operating in NAFO area.

Age reading error has two major elements: bias and precision. The **bias** of age readings is caused by a consistent deviation of reading results between readers and is skewed from the mean to one side or the other, while **precision** of age readings measures the closeness of repeated independent age estimates. Precision reflects the degree of agreement among readers and is not to be confused with accuracy, which relates to the agreement with the true age of the fish (Wilson *et al*., 1987, ICES 1996, in Stransky *et al*., 2002).

This Greenland halibut otoliths exchange was carried out in order to examine for differences in age interpretations among age readers from Spain, Canada and Portugal.

Material and Methods

A total of 100 otoliths mainly covering the length range of Greenland halibut caught in the commercial fishery and Flemish Cap Survey (NAFO Div. 3LM-NO) were sent from Spain to Canada, and then to Portugal. The length and ageing material were collected during 2001. Both otoliths per fish are removed and stored dry. Three different methods are in use in the different Institutes involved in this exc hange.

In Spain, Esther Roman (**ESP**) read the otoliths. Both otoliths have been stored in small vials and soaked in glycerine-alcohol (10:90) for approximately 12 hours prior to examination, which made the hyaline-opaque zones more distinct. Left/right otoliths were examined under binocular microscope with reflected light illuminator inclined at 45-60 degrees at magnifications of 10-16x. Under reflected light, winter zones (hyaline) appear dark and summer (opaque) zones are white. Otoliths are placed in black plastic or clear glass holders with a black base or background and then immersed in the wetting agent (glycerine-alcohol) for viewing. The ages were assigned by observing the prepared otoliths. Interpretation of age was some times difficult, especially for fish over 10 years old.

In Canada, two St. John's Greenland halibut age readers (Brian Greene - **CAN1** and Randy Burry - **CAN2**) read the otoliths. The methodology was the same for each reader but each set of readings was done independently. Untreated otoliths were placed in a black watch glass (small bowl) and were covered with 95% ethanol. The otoliths were then read at approximately 10x magnification in reflected light, convex side up. For this exchange Canada decided not to grind any of the otoliths so as not to interfere with any future readers.

In Portugal, Ricardo Alpoim read the otoliths. Two methods was used to prepare the otoliths:

PRT1 - the otoliths were immersed in immersion oil for 24 hours and then read.

PRT2 - the otoliths were soaked in glycerine-thymol (50:50) for approximately 72 hours, and after that burned in a oven for 30 min at 200ºC (Godinho and Alpoim, 1997), then were immersed in immersion oil for 24 hours and then read. This method is the one used regularly in Portugal coming from the Reyjavik workshop 1996.

In both methods the otoliths were read at approximately 10x magnification in reflected polarised light, in most otoliths convex side up, but for the bigger ones changing between the convex and the concave size up was needed. The second lecture was made one week after the first, without looking to the first results.

For the comparison of bias and precision between readers a set of statistical tests and graphical methods were applied.

Simple linear regression analysis, the parametric paired *t*-test and the nonparametric Wilcoxon matched-pairs rank test were used for detect bias. The slope and intercept of simple linear regressions are tested for significant differences ($\alpha = 0.05$) from 1.0 and 0, respectively. The parametric paired *t*-test and the nonparametric Wilcoxon matched pairs rank test are used to detect significant differences from a paired difference of 0. Error terms are 95% confidence limits. Age bias plots are presented to visual detect the systematic age differences between readers, Each error bar represents the 95% confidence interval about the mean age assigned by one ager for all fish assigned a given age by a second ager. This allowed the detection of linear and nonlinear bias (Campana *et al.,* 1995).

The traditional index of precision in ageing studies, percent agreement, is gradually falling out of favor, because percent agreement may vary substantially among species and within a species. Beamish and Fournier (1981) illustrated this point by nothing that 95% agreement of two readers to within 1 year can be poor precision for ageing Pacific cod (*Gadus macrocephalus*), given the few year-classes in the fishery, but 95% agreement to within 5 years can be good precision for spiny dogfish (*Squalus acanthias*), given this species' 60year longevity. Beamish and Fournier (1981), therefore, suggested an average percent error, which is dependent on the average age of the fish species observed. Chang (1982) modified this index to a coefficient of variation, substituting the absolute deviation by the standard deviation from the mean age (Campana *et al*., 1995). Besides these indices, the correlation coefficient *r* is given to evaluate the fraction of variation explained by the linear relationship between readers or otolith preparation methods.

Results

Table 1 and panels A and B from Fig. 1 to 10 presents the results of the statistical tests for detection of bias between readers.

The regression analysis shows some bias between all readers with the exception of CAN/PRT1 and CAN1/CAN2. This bias is more accentuated between ESP and CAN. Between CAN/PRT1, CAN1/CAN2 and PRT1/PRT2 the intercept gives close to zero.

By the paired *t*-test with 95% significance levels, between all the readers a mean pair difference from 0.21 to 0.69 years old is observed, with the exception between ESP and CAN that the mean pair difference is close to zero.

The Wilcoxon test shows the same results as the paired *t*-test.

From the plots bias is detected also between some readers been the main difference that CAN give less age than ESP and PRT to the older fish (around >9 years old).

Table 2 and panels C and D from Fig. 1 to 10 presents the results of measures of precision between readers.

The correlation coefficient is relatively high between readers, but the percent of agreement (tolerance +/- 0 years) is low (from 37 to 52% with the exception of PRT1/PRT2).

The other two measures, the coefficient of variation (CV) and the average percent error (APE), gives apparent high levels of precision between all the readers however precision is lower between ESP and CAN and higher between PRT1 and PRT2.

From the plots the percent of agreement decrease between readers with the ageing of older fish. If the tolerance is set in +/- 2 years old the percent of agreement between readers increase to levels higher than 90%. The CV by age is relative stable along the age range, showing sometimes less precision between readers in the younger or older ages.

Conclusions

Although most of the differences in age interpretations were within one to two years, there were strong biases in several cases.

The Spanish age reader had a tendency to assign lower ages to fish younger than 9 years old and higher ages to the older ones, relative to the Canadian age readers. This can be observed in the age-bias plots and in the regression but the t-test or the Wilcoxon test does not detect it.

The age reader from Portugal, like the Spanish age reader, assigned higher ages to the older fish (>8 years old) than Canada, but for the younger fish the agreement was very good.

Burning the otoliths before reading seems to enhance the last few annuli from the older fish.

References

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

CAMPANA, S. E., M. C. ANNAND, and J. I. McMILLAN. 1995. Graphical and statistical methods for determining the consistency of age determinations. *Trans. Am. Fish. Soc.*, 124: 131-138.

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

CONOVER, W. J. 1998. *Practical nonparametric statistics.* 3rd ed., John Wiley & Sons, New York, 584 pp.

GODINHO, L. and R. Alpoim, 1997**.** Essays on baking tecnique of Greenland halibut otoliths for ageing purposes. *NAFO SCR Doc.* 97/67, Ser. No N2901, 3pp.

ICES 1996. Report of the workshop on age reading of *Sebastes* spp. *ICES C.M.* 1996/G:1, 32 pp.

STRANSKY, C., S. GUÐMUNDSDÓTTIR, Þ. SIGURÐSSON, SV. LEMVIG, and K. NEDREAAS. 2002. Age readings of *Sebastes mentella* otoliths: bias and precision between readers. NAFO SCR Doc 02/26, Serial nº N4632, 11 p.

WILSON, C. W., R. J. BEAMISH, K. D. BROTHERS, K. D. CARLANDER, J. M. CASSELMAN, J. M. DEAN, A. JEARLD Jr., E. D. PRINCE, and A. WILD. 1987. Glossary. *In*: Age and growth of fish. Summerfelt, R.C. and G.E. Hall, G.E. (eds.), Iowa State University Press, Ames, Iowa, pp. 527-530.

Tab. 1: Statistical tests for the detection of bias between readers.

Statistic	CAN1 versus ESP (N = 100)	CAN2 versus ESP (N = 100)	PRT1 versus ESP (N = 100)	PRT2 versus ESP (N = 100)	CAN2 versus CAN1 (N = 100)	PRT1 versus CAN1 (N = 100)	PRT2 versus CAN1 (N = 100)	PRT1 versus CAN2 (N = 100)	PRT2 versus CAN2 (N = 100)	PRT2 versus PRT1 (N = 100)
Regression										
Slope	0.76 ± 0.05	0.78 ± 0.05	0.87 ± 0.04	0.93 ± 0.04	0.96 ± 0.06	1.04 ± 0.07	1.11 ± 0.06	1.02 ± 0.06	1.09 ± 0.06	1.04 ± 0.03
P	0.000	0.000	0.000	0.002	0.301	0.233	0.002	0.424	0.004	0.008
Intercept	1.51 ± 0.48	1.60 ± 0.48	1.16 ± 0.38	0.96 ± 0.37	0.45 ± 0.50	0.14 ± 0.57	-0.1 ± 0.56	0.03 ± 0.57	-0.2 ± 0.54	-0.1 ± 0.28
P	0.000	0.000	0.000	0.000	0.079	0.630	0.583	0.896	0.331	0.385
Paired t-test										
Mean paired difference	-0.25 ± 0.28	-0.04 ± 0.27	0.21 ± 0.19	0.44 ± 0.17	0.21 ± 0.20	0.46 ± 0.22	0.69 ± 0.23	0.25 ± 0.22	0.48 ± 0.22	0.23 ± 0.12
P	0.087	0.776	0.042	0.000	0.044	0.000	0.000	0.031	0.000	0.000
Wilcoxon test										
Positive ranks	31	34	40	43	36	36	43	28	39	24
Negative ranks	32	29	16	8	20	12	8	21	14	5
Ties	37	37	44	49	44	52	49	51	47	71
P	0.122	0.943	0.034	0.000	0.083	0.000	0.000	0.046	0.000	0.001

Tab. 2: Measures of precision between readers:

Statistic or index	CAN1 versus ESP (N = 100)	CAN2 versus ESP (N = 100)	PRT1 versus ESP (N = 100)	PRT2 versus ESP (N = 100)	CAN2 versus CAN1 (N = 100)	PRT1 versus CAN1 (N = 100)	PRT2 versus CAN1 (N = 100)	PRT1 versus CAN2 (N = 100)	PRT2 versus CAN2 (N = 100)	PRT2 versus PRT1 (N = 100)
Correlation coefficient (r)	0.938	0.941	0.968	0.973	0.950	0.946	0.954	0.948	0.957	0.987
Coefficient of variation (%) [a]	9.13	9.58	7.21	7.03	6.45	5.81	6.34	6.19	6.38	2.61
Average percent error [b]	6.46	6.78	5.10	4.97	4.56	4.11	4.48	4.37	4.51	1.84
Percent agreement	37.0	37.0	44.0	49.0	44.0	52.0	49.0	51.0	47.0	71.0

[a] from Chang (1982)
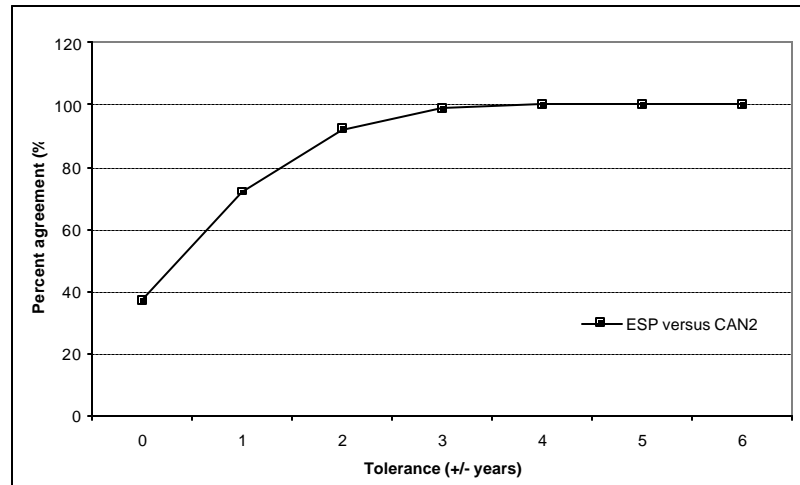
[b] from Beamish and Fournier (1981)
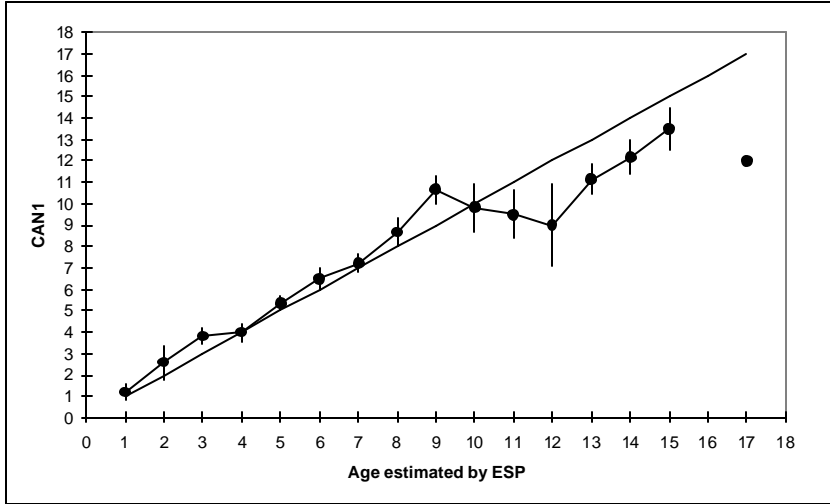
A - Age bias graphs. Each error bar represents the 95% confidence interval about the mean age assigned by one reader for all fish assigned a given age by a second ager. The 1:1 equivalence (solid line) is also indicated.
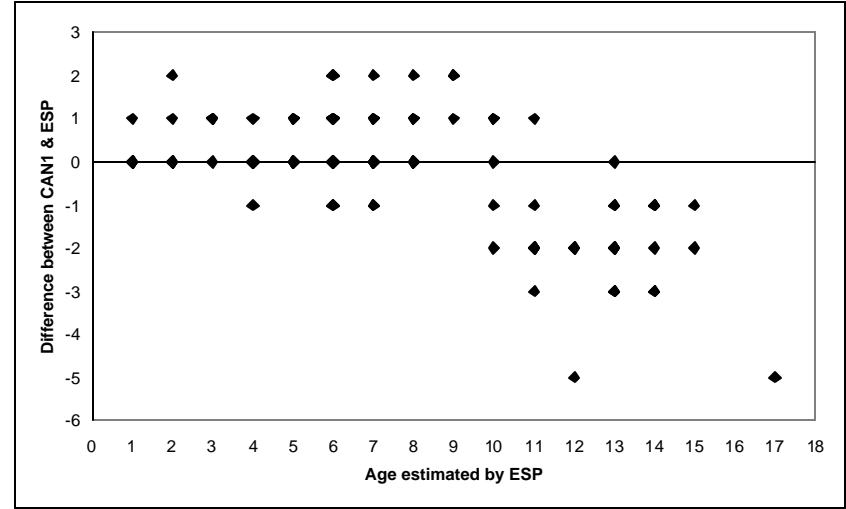
B - Differences (years) in age estimates between paired age readers. Each point represents one or more observations of an individual fish.
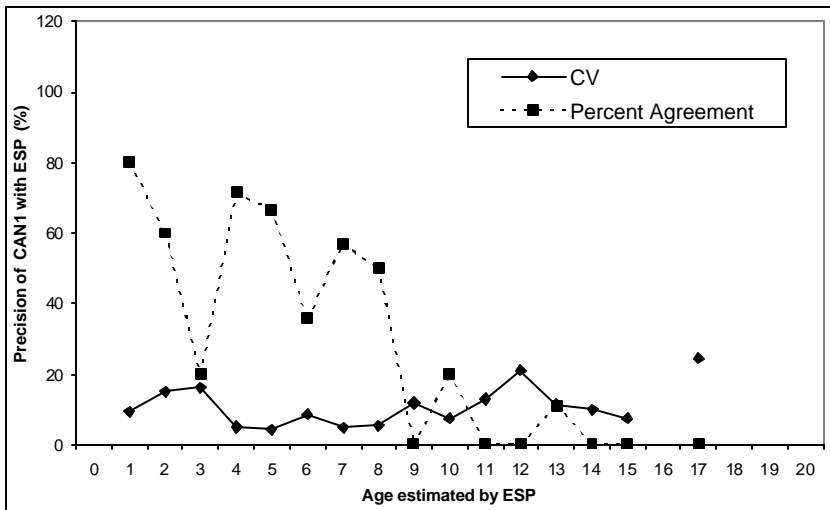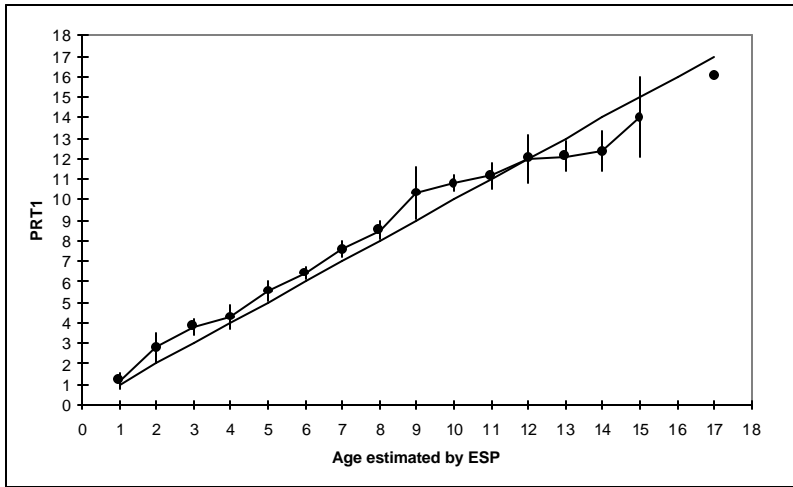
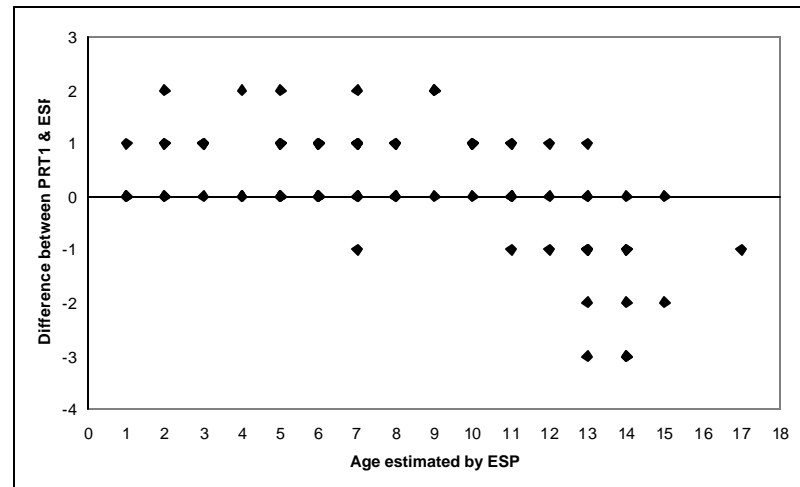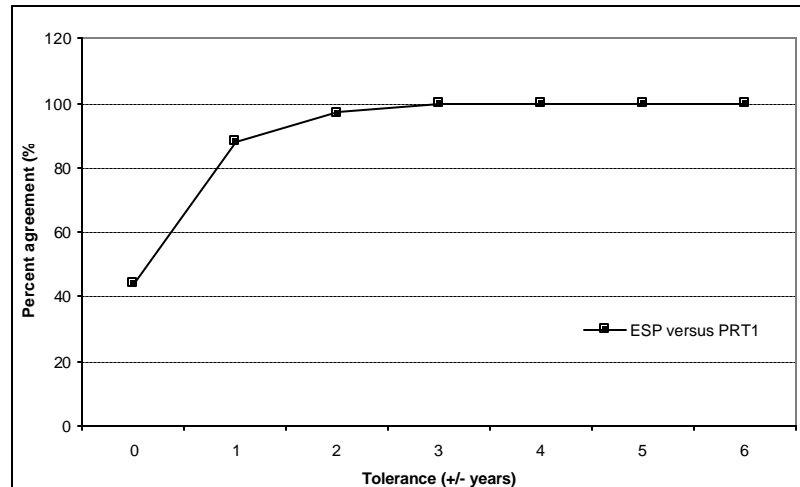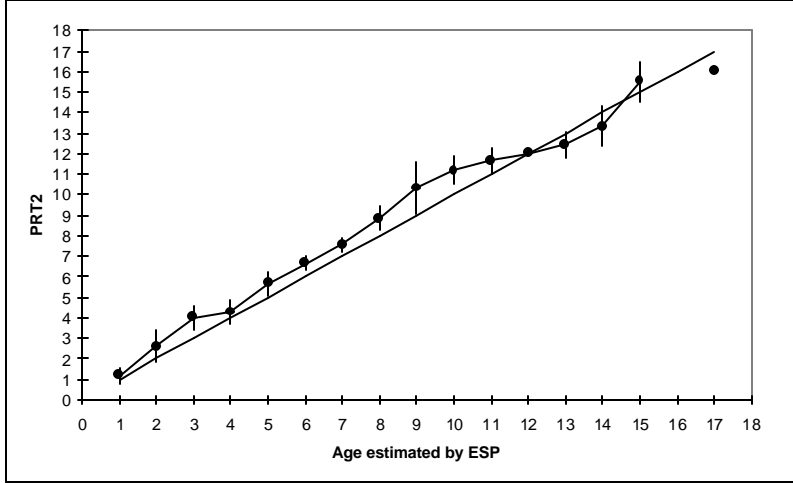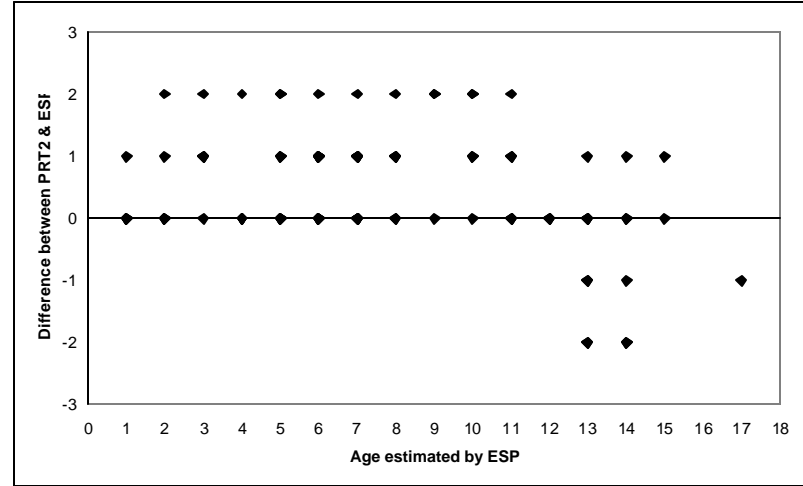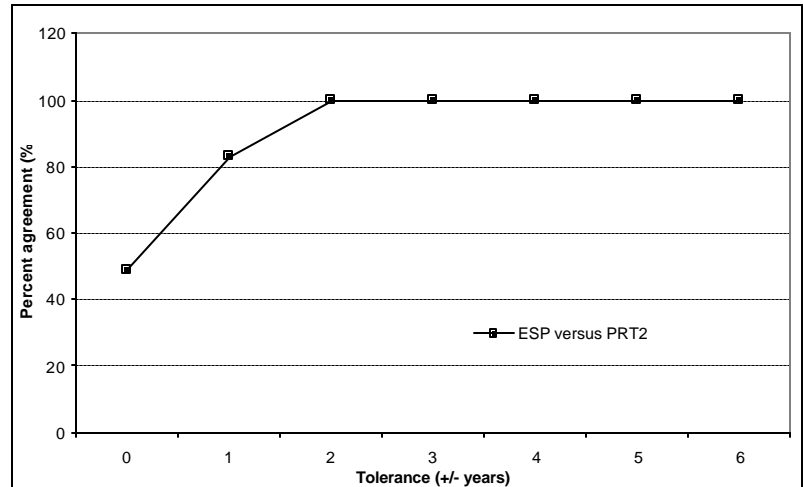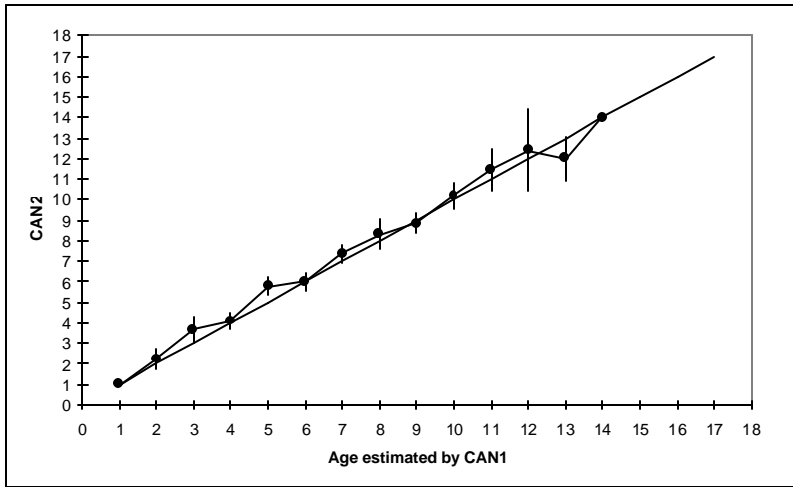C - Coefficient of variation (CV) and percent agreement for two of the pairwise age comparisons.
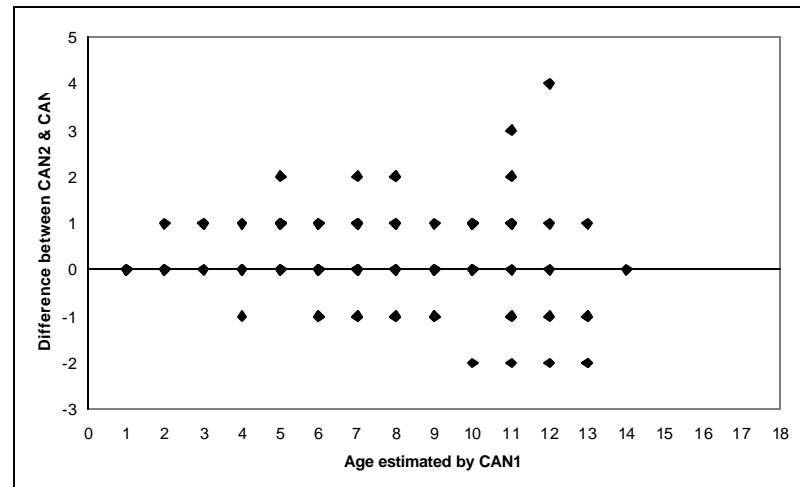
D - Percent agreement for reader comparisions for a tolerance level of +/- 0 to +/- 6 years applied to all age groups assigned by ESP.

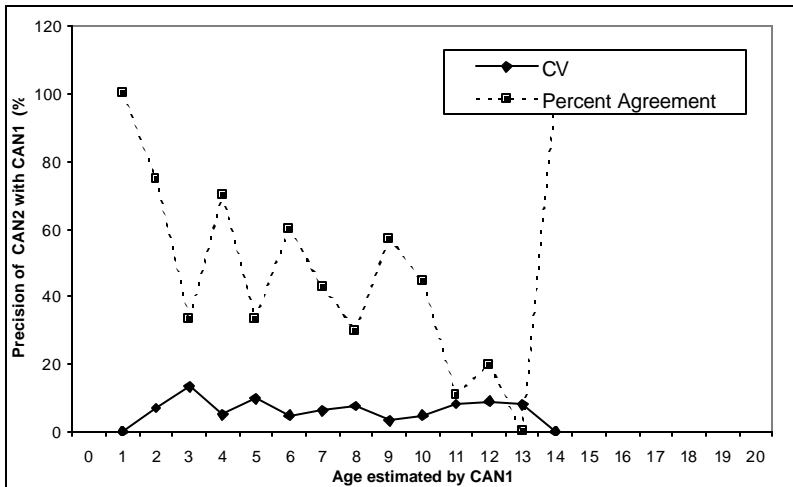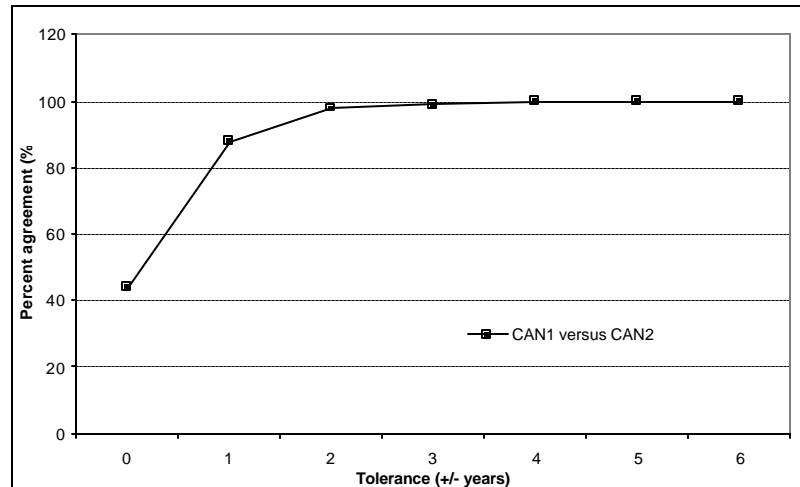**Fig. 1: Graphical methods for comparison between readers: ESP and CAN1**
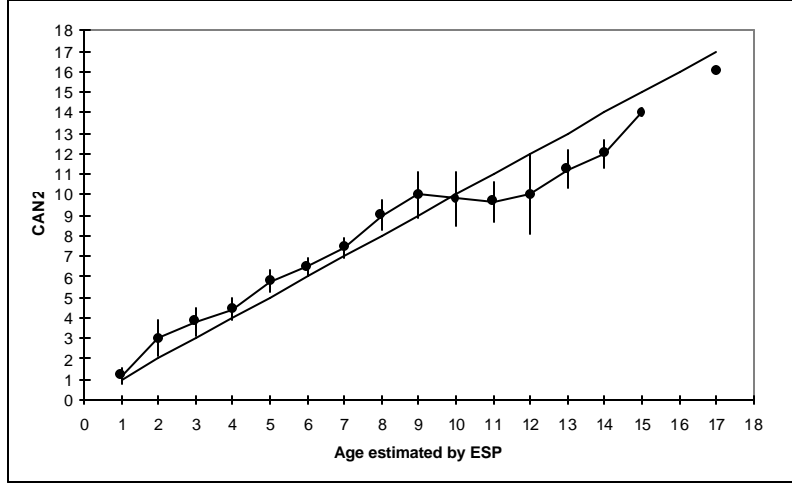
A - Age bias graphs. Each error bar represents the 95% confidence interval about the mean age assigned by one reader for all fish assigned a given age by a second ager. The 1:1 equivalence (solid line) is also indicated.
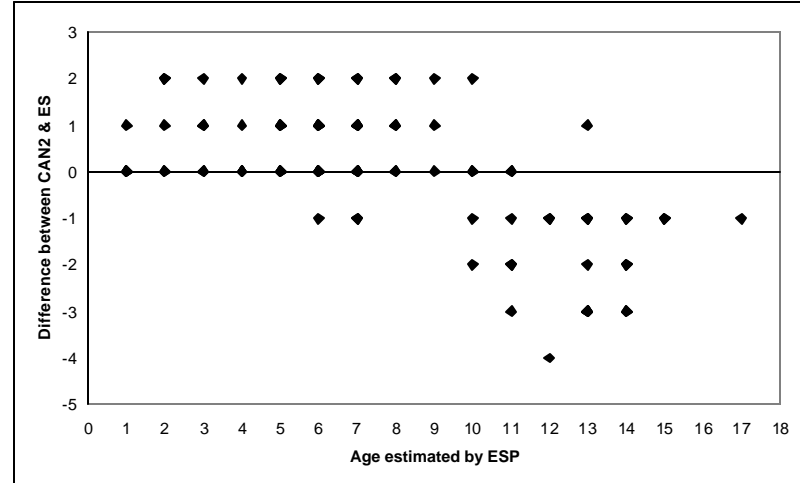
B - Differences (years) in age estimates between paired age readers. Each point represents one or more observations of an individual fish.
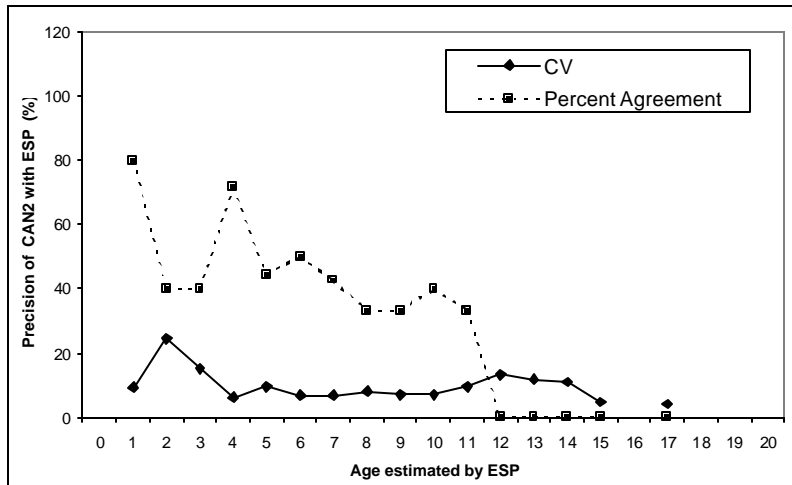
C - Coefficient of variation (CV) and percent agreement for two of the pairwise age comparisons.

D - Percent agreement for reader comparisions for a tolerance level of +/- 0 to +/- 6 years applied to all age groups assigned by ESP.

Fig. 2: Graphical methods for comparison between readers: ESP and CAN2

A - Age bias graphs. Each error bar represents the 95% confidence interval about the mean age assigned by one reader for all fish assigned a given age by a second ager. The 1:1 equivalence (solid line) is also indicated.

B - Differences (years) in age estimates between paired age readers. Each point represents one or more observations of an individual fish.
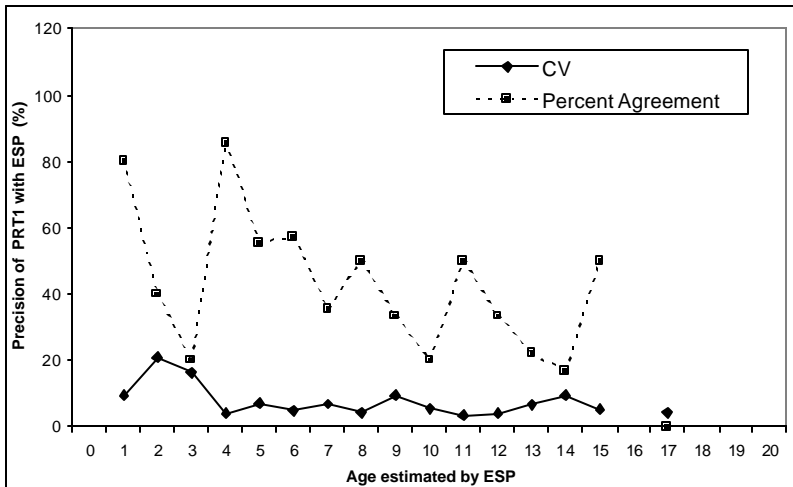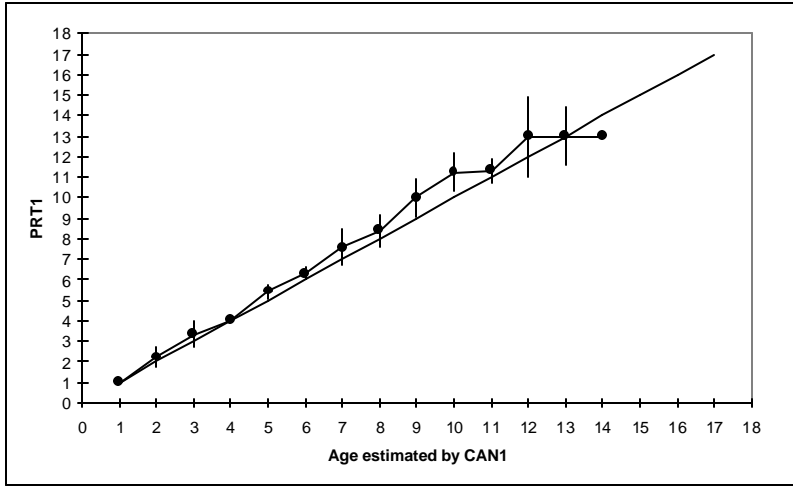
C - Coefficient of variation (CV) and percent agreement for two of the pairwise age comparisons.
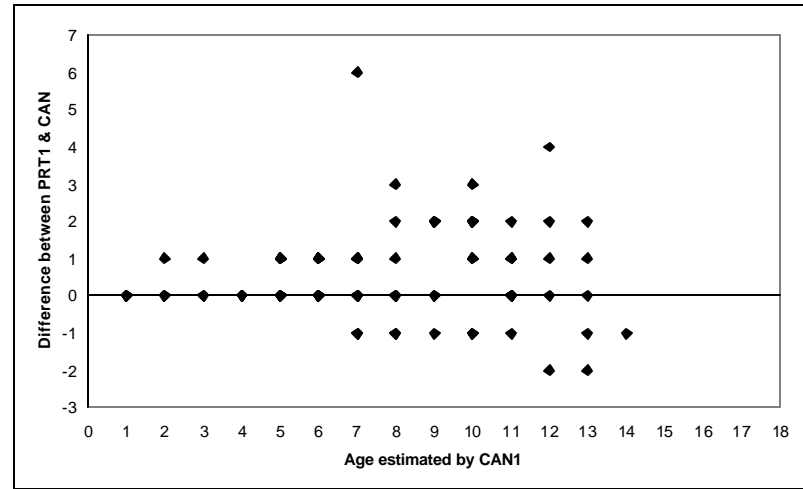
D - Percent agreement for reader comparisions for a tolerance level of +/- 0 to +/- 6 years applied to all age groups assigned by ESP.
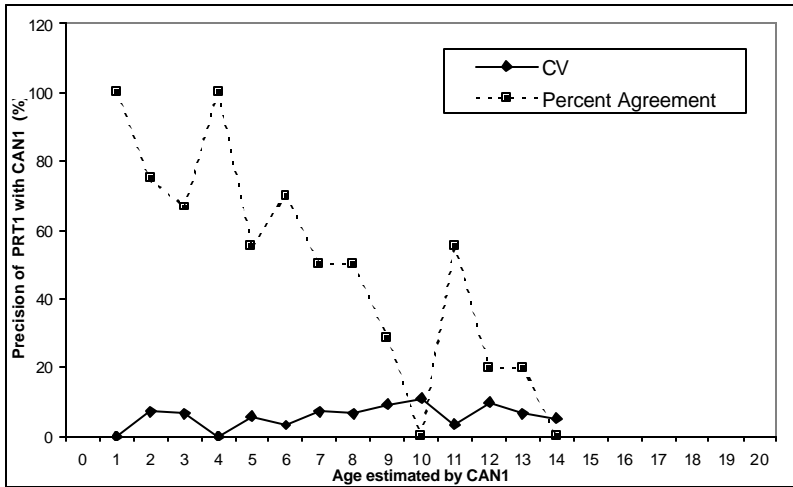
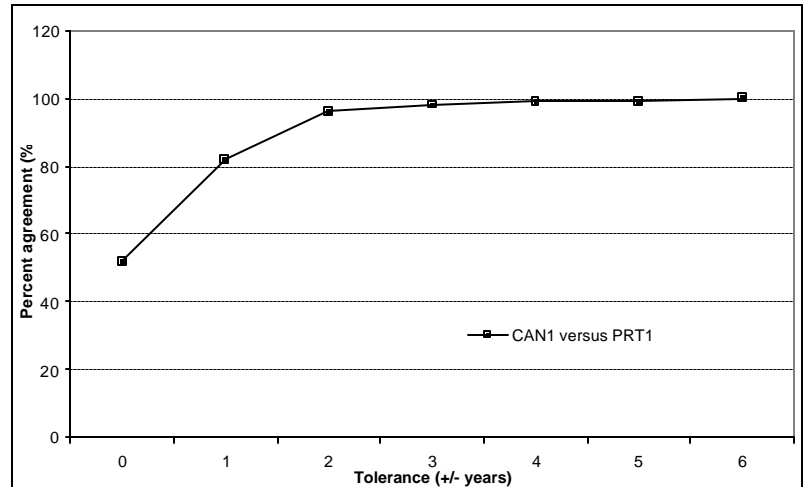**Fig. 3: Graphical methods for comparison between readers: ESP and PRT1**

A - Age bias graphs. Each error bar represents the 95% confidence interval about the mean age assigned by one reader for all fish assigned a given age by a second ager. The 1:1 equivalence (solid line) is also indicated.

B - Differences (years) in age estimates between paired age readers. Each point represents one or more observations of an individual fish.
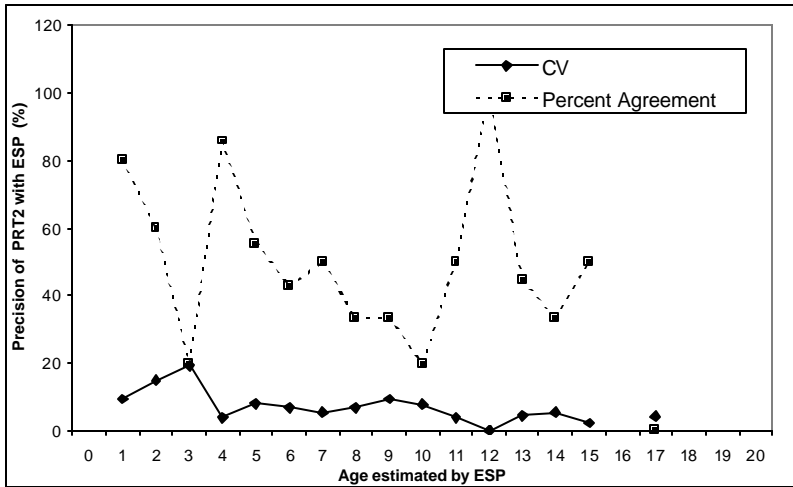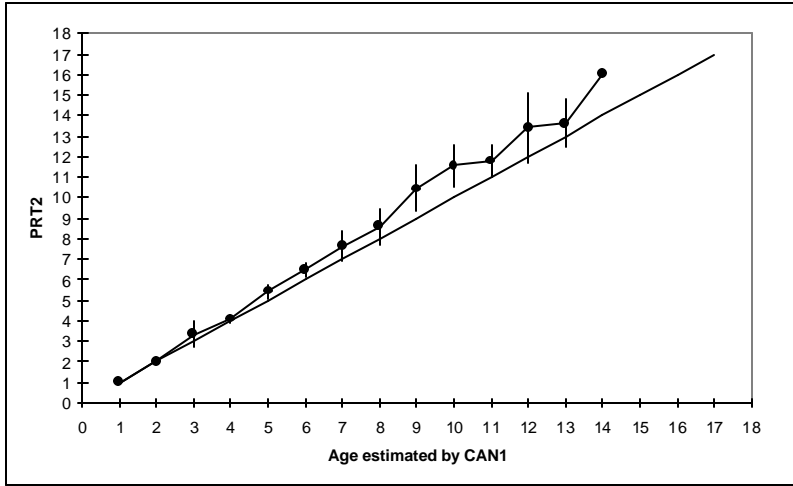
C - Coefficient of variation (CV) and percent agreement for two of the pairwise age comparisons.
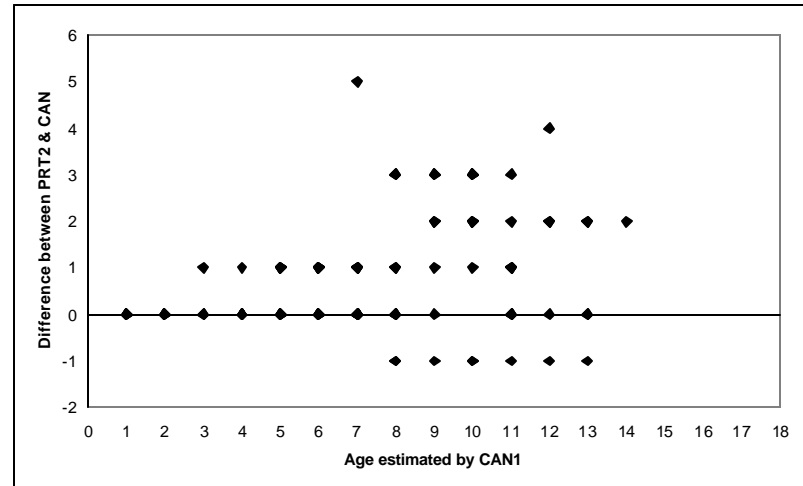
D - Percent agreement for reader comparisionsfor a tolerance level of +/- 0 to +/- 6 years applied to all age groups assigned by ESP.

Fig. 4: Graphical methods for comparison between readers: ESP and PRT2

A - Age bias graphs. Each error bar represents the 95% confidence interval about the mean age assigned by one reader for all fish assigned a given age by a second ager. The 1:1 equivalence (solid line) is also indicated.

B - Differences (years) in age estimates between paired age readers. Each point represents one or more observations of an individual fish.

C - Coefficient of variation (CV) and percent agreement for two of the pairwise age comparisons.

D - Percent agreement for reader comparisionsfor a tolerance level of +/- 0 to +/- 6 years applied to all age groups assigned by CAN1.
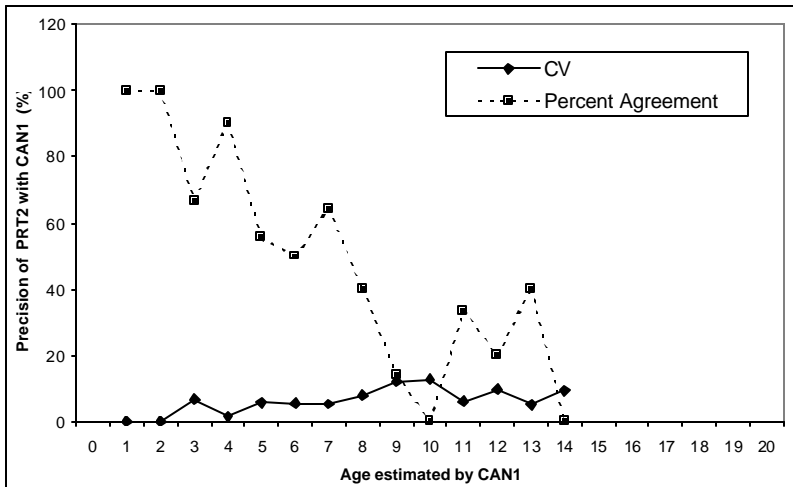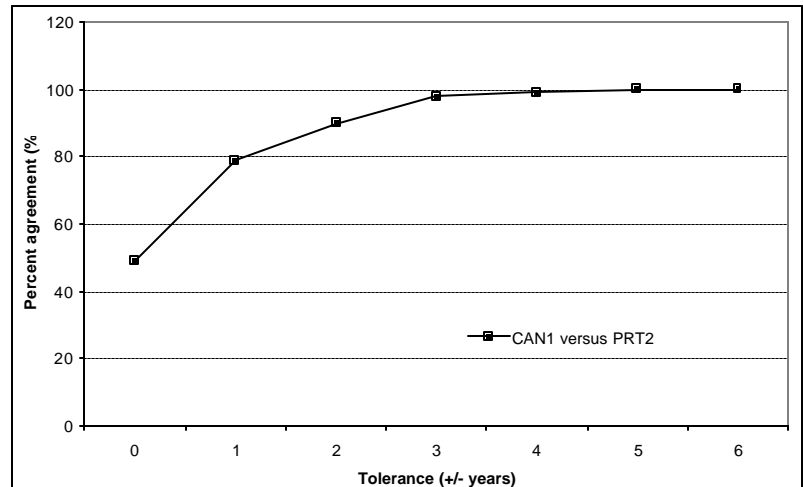
Fig. 5: Graphical methods for comparison between readers: CAN1 and CAN2

A - Age bias graphs. Each error bar represents the 95% confidence interval about the mean age assigned by one reader for all fish assigned a given age by a second ager. The 1:1 equivalence (solid line) is also indicated.

B - Differences (years) in age estimates between paired age readers. Each point represents one or more observations of an individual fish.
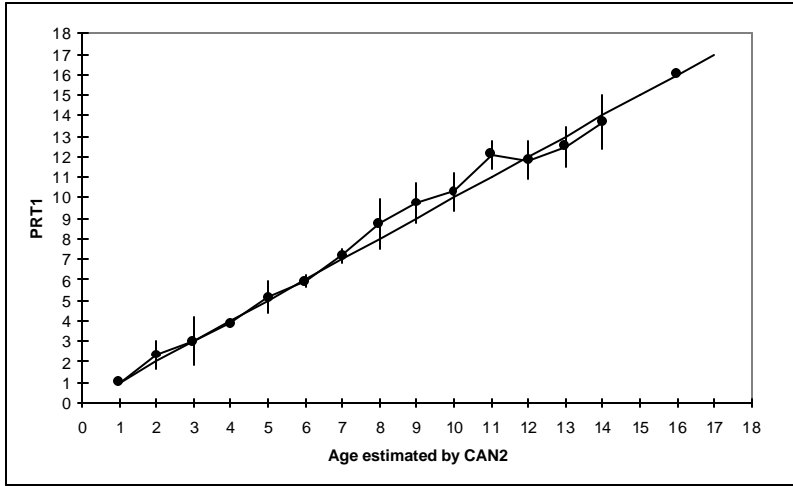
C - Coefficient of variation (CV) and percent agreement for two of the pairwise age comparisons.
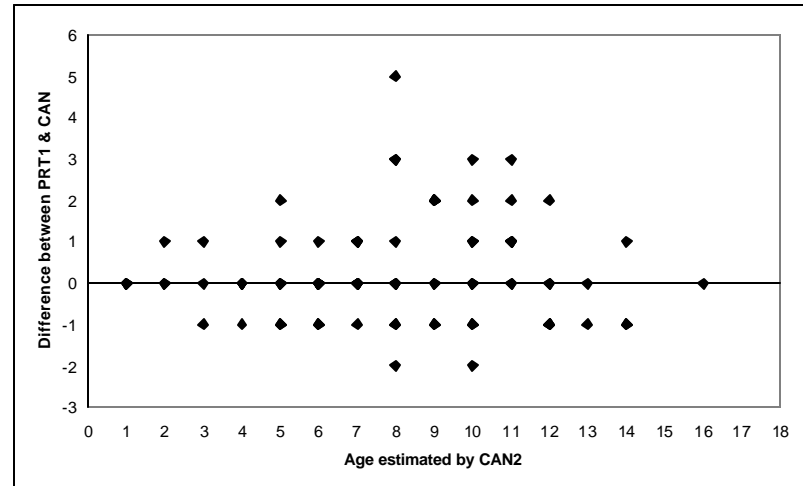
D - Percent agreement for reader comparisions for a tolerance level of +/- 0 to +/- 6 years applied to all age groups assigned by CAN1.
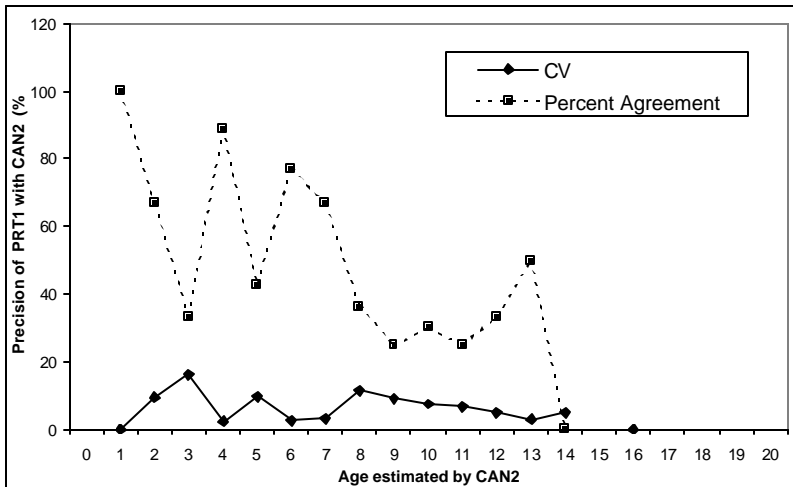
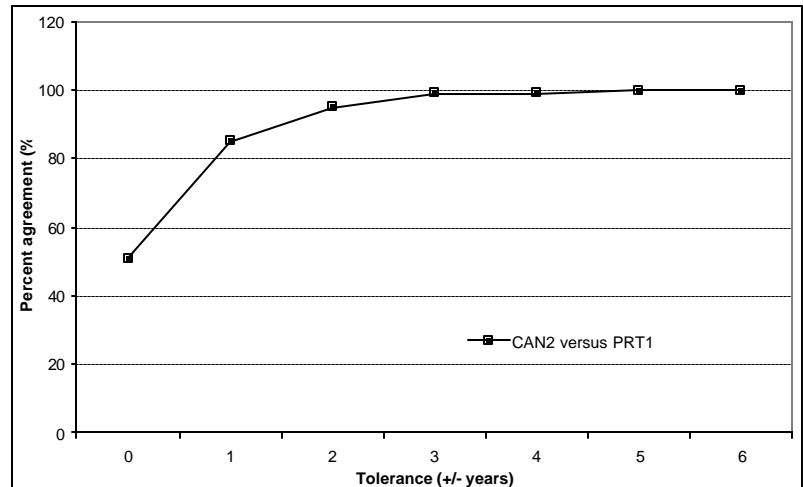**Fig. 6: Graphical methods for comparison between readers: CAN1 and PRT1**

A - Age bias graphs. Each error bar represents the 95% confidence interval about the mean age assigned by one reader for all fish assigned a given age by a second ager. The 1:1 equivalence (solid line) is also indicated.

B - Differences (years) in age estimates between paired age readers. Each point represents one or more observations of an individual fish.
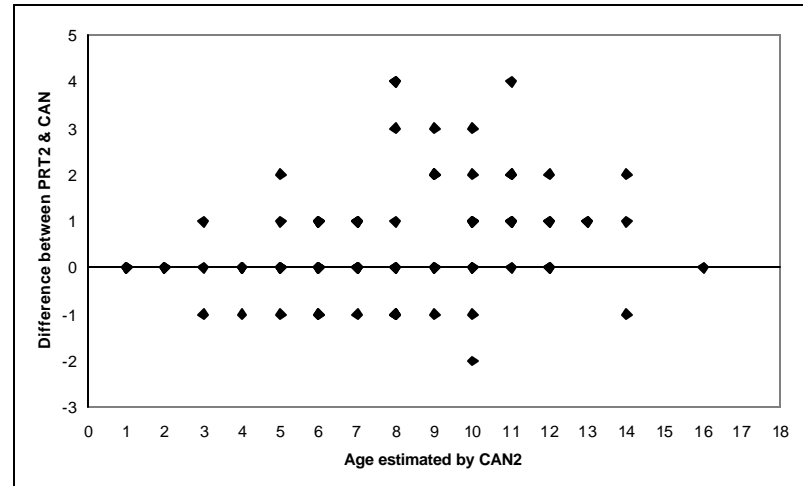
C - Coefficient of variation (CV) and percent agreement for two of the pairwise age comparisons.

D - Percent agreement for reader comparisionsfor a tolerance level of +/- 0 to +/- 6 years applied to all age groups assigned by CAN1.

Fig. 7: Graphical methods for comparison between readers: CAN1 and PRT2

A - Age bias graphs. Each error bar represents the 95% confidence interval about the mean age assigned by one reader for all fish assigned a given age by a second ager. The 1:1 equivalence (solid line) is also indicated.

B - Differences (years) in age estimates between paired age readers. Each point represents one or more observations of an individual fish.
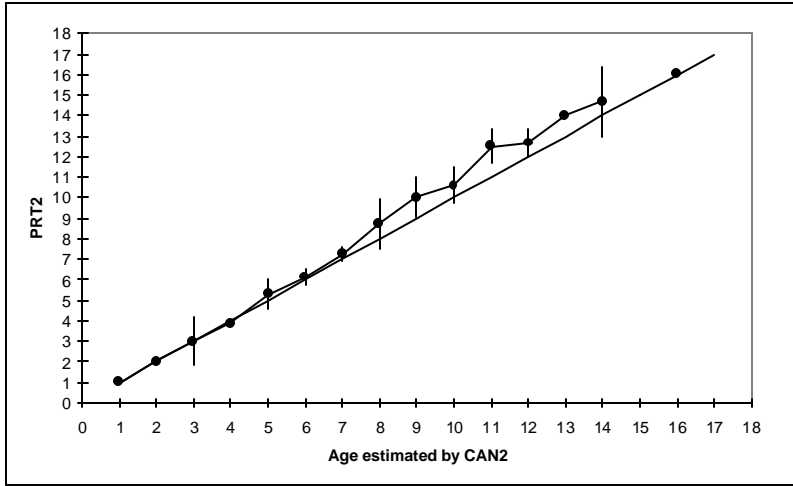
C - Coefficient of variation (CV) and percent agreement for two of the pairwise age comparisons.
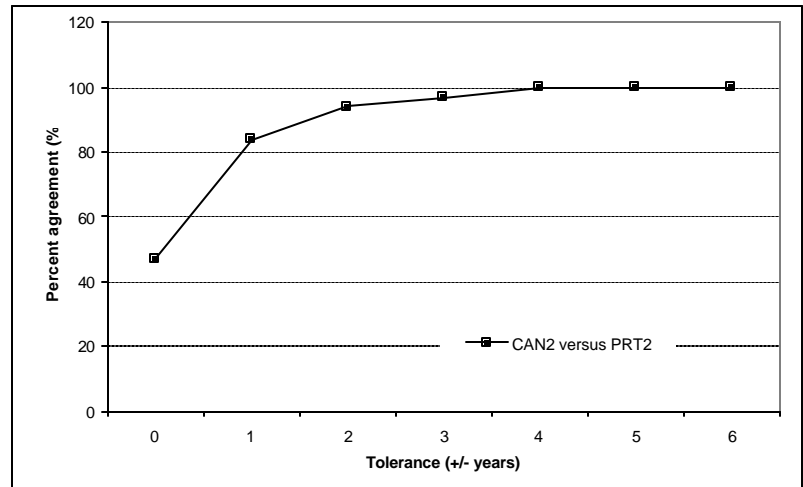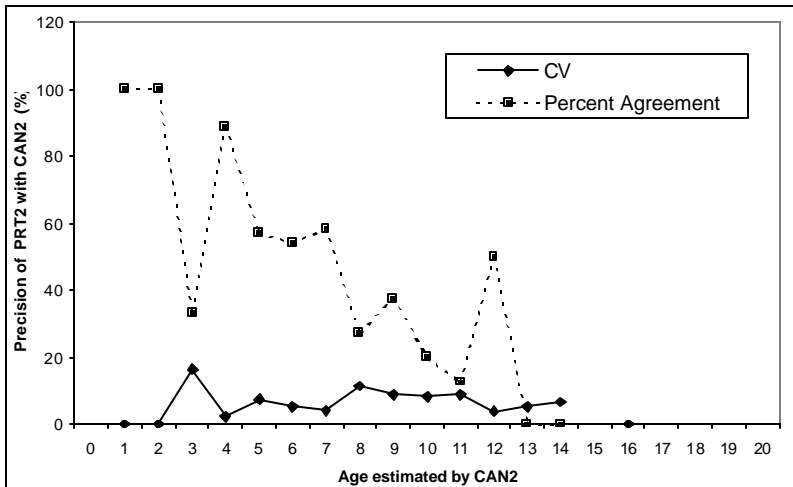
D - Percent agreement for reader comparisionsfor a tolerance level of +/- 0 to +/- 6 years applied to all age groups assigned by CAN2.

Fig. 8: Graphical methods for comparison between readers: CAN2 and PRT1

A - Age bias graphs. Each error bar represents the 95% confidence interval about the mean age assigned by one reader for all fish assigned a given age by a second ager. The 1:1 equivalence (solid line) is also indicated.
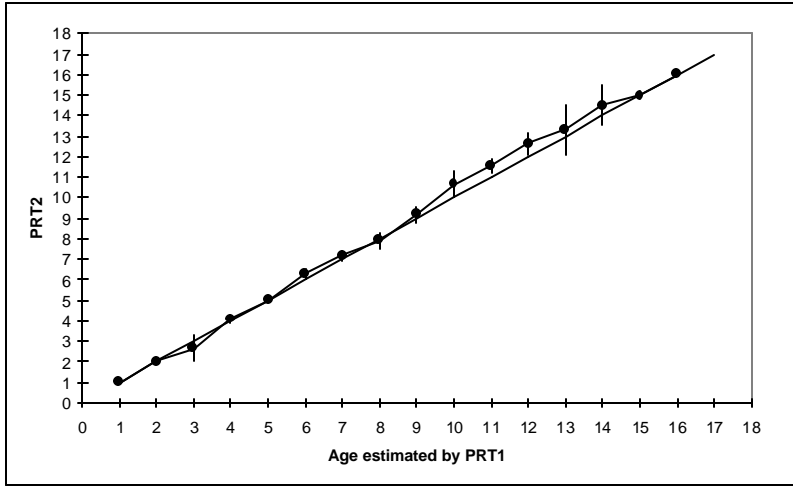
B - Differences (years) in age estimates between paired age readers. Each point represents one or more observations of an individual fish.
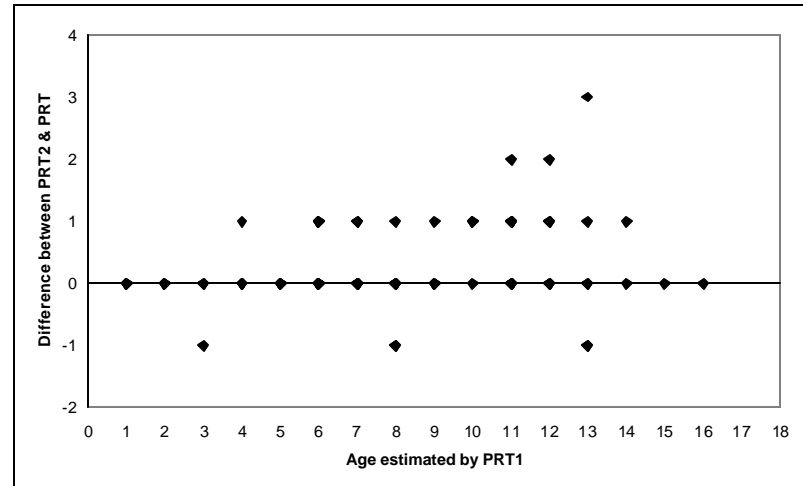
C - Coefficient of variation (CV) and percent agreement for two of the pairwise age comparisons.

D - Percent agreement for reader comparisionsfor a tolerance level of +/- 0 to +/- 6 years applied to all age groups assigned by CAN2.
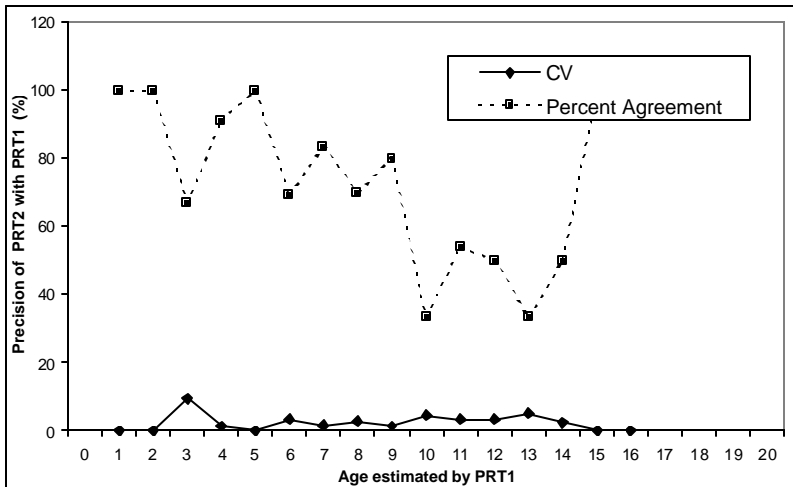
Fig. 9: Graphical methods for comparison between readers: CAN2 and PRT2
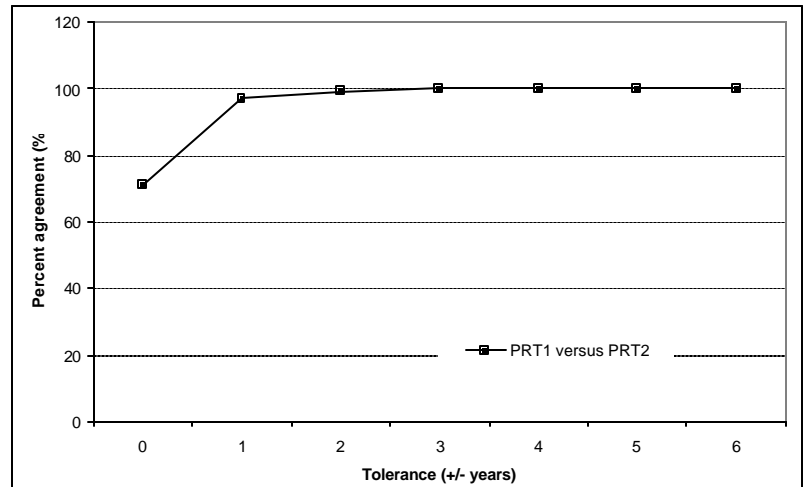
A - Age bias graphs. Each error bar represents the 95% confidence interval about the mean age assigned by one reader for all fish assigned a given age by a second ager. The 1:1 equivalence (solid line) is also indicated.

B - Differences (years) in age estimates between paired age readers. Each point represents one or more observations of an individual fish.

C - Coefficient of variation (CV) and percent agreement for two of the pairwise age comparisons.

D - Percent agreement for reader comparisionsfor a tolerance level of +/- 0 to +/- 6 years applied to all age groups assigned by PRT1.

Fig. 10: Graphical methods for comparison between readers: PRT1 and PRT2