SC/62/IA2

Examination of the period/reader effect on the age-determination for the Antarctic minke whales and its implications to the statistical catch-at-age analyses

TOSHIHIDE KITAKADO[1] AND ANDRÉ E. PUNT[2]

[1] *Tokyo University of Marine Science and Technology, 5-7, Konan 4, Minato-ku, Tokyo 108-8477, Japan*
[2] *School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195-5020, USA*
Contact email: kitakado@kaiyodai.ac.jp

ABSTRACT

The impact of period/reader on age-determination by three Japanese readers is explored by comparing estimates of age from earplugs from a control reader with age-estimates by the Japanese readers. A total of 250 plugs selected according to a predetermined protocol were used in the analyses. Parameters determining ageing error matrices were estimated using a maximum likelihood method. The results demonstrated that the Japanese readers and the control reader differed in terms of both expected age given true age and variance in age-estimates. The results also suggested that the expected age and random uncertainty in age-estimates differed among the Japanese readers. This work could contribute to how catch-at-age data are used in the statistical catch-at-age analyses and in future virtual population analyses.

KEYWORDS: ANTARCTIC MINKE WHALES; AGE-READING ERROR; EARPLUGS

INTRODUCTION

Estimation methods based on age-structured population dynamics models are common tools for investigating the status of cetacean populations, and these methods have been advanced recently. They can utilize catch-at-age data, which are created by dedicated work by age-readers, for parameter estimation purposes. Ageing is therefore one of key techniques for improving age-structured analyses. This is the case for Antarctic minke whales (e.g., Mori *et al*. 2007, Punt and Polacheck 2007).

Catch-at-age data for Antarctic minke whales sampled by commercial and scientific whaling are now available for many years due to a considerable amount of age-reading work. On the other hand, there is also concern regarding ageing bias. Specifically, the IWC/SC has recognized some inconsistency in the length-at-age data between the commercial and JARPA catches and it identified that ageing error may be one of possible causes for this inconsistency (IWC 2009).

The IWC/SC therefore proposed an age-reading experiment to determine if there are systematic differences in ageing error among Japanese age-readers. It also suggested that an ageing experiment could be undertaken to evaluate such potential differences (IWC 2009). The sample size of 250 animals was proposed and justified by Kitakado (2009) through a simulation study (see also IWC 2009). The IWC/SC appointed Christina Lockyer as a control reader. She conducted the ageing experiment in Japan and provided ageing outcomes (estimates of age) along with her age-reading protocol, which were kept consistent throughout the experiment (Lockyer 2010). The data obtained by Lockyer were compared with the age-estimates from the Japanese readers.

The aim of this paper is to investigate the extent of ageing error for the Antarctic minke whales by comparing age-estimates from the two groups: Lockyer and the Japanese scientists. The results of this work could contribute to how catch-at-age data are used in the statistical catch-at-age analyses (Punt 2010) and in future virtual population analyses.

MATERIALS AND METHODS

Materials

A total of 250 plugs were employed in this analysis (see Table 1). These plugs had already been read by the three Japanese scientists; Masaki and Kato read each plug once while Zenitani read each plug three times and then selected a best estimate based on the three estimates. 50 plugs were randomly chosen for each of five periods (Table 1) and read twice by the independent reader (Lockyer). Ten of the 50 plugs for each period were also randomly selected and read by Lockyer three times. The readings were stratified by period to assess whether changes over time occurred in ageing bias and random ageing error. It should be noted that the 250 samples were restricted to those which could be read using the Japanese protocol.

Lockyer (2010) was unable to read all of the plugs (Table 2). Age estimates could be obtained for more than 86% of the plugs for each trial, although the proportion of plugs which could be read decreased between the first and second trials. The bulk of the analyses are based on the "valid" readings only, although sensitivity tests consider the use of the data from the other categories in Table 2 (see below).

Statistical model

Suppose that two groups of readers independently obtain age-estimates using a common set of n samples (here n =250). Group 1 consists of only one reader (Lockyer) who conducted ageing at most three times for the n samples (Group1 = Reader 1 hereafter). Group 2 consists of three readers (Masaki, Kato and Zenitani) and they read different plugs from different time periods. The sample sizes of their readings are n_M =50 and n_K =78 (Table 1). Only one Japanese reader (Zenitani) read plugs multiple times (n_Z =122).

Let $a_{1jk}\ (j=1,2,...,n; k=1,...,r_j)$ be the observed age by Group 1 (assuming that it is a "valid" count) of the j-th sample during the k-th of r_j trials ($r_j = 2 \text{ or } 3$). Similarly, for Group 2, let $a_{2j}\ (j=1,2,...,n_M)$ and $a_{2j}\ (j=n_M+1,2,...,n_M+n_K)$ respectively denote observed ages by Masaki and Kato and $a_{2jk}\ (j=n_M+n_K+1,2,...,n; k=1,2,3)$ denote observed counts for the j-th sample during the k-th trial by Zenitani. As mentioned earlier, Lockyer did not assign "valid" ages to all of the samples during all of the trials (Table 2). In such cases, the notation changes accordingly. For example, the data for the "either" or "interval" categories can be denoted $a^{(1)}_{1jk}$ and $a^{(2)}_{2jk}$ respectively (for "either" the age is either $a^{(1)}_{1jk}$ or $a^{(2)}_{2jk}$ while for "interval" the age-estimate is between $a^{(1)}_{1jk}$ and $a^{(2)}_{2jk}$).

Now, consider the joint probability distribution of the observations. Let $b_i(a;\phi)$ and $\sigma_i(a;\phi)$, respectively, denote the expected age and standard deviation for the age-estimates for the i-th Group for an animal of true age a, where ϕ represents a vector of unknown parameters. The variability in ageing is expressed as a matrix form $\{P_i(a'|a;\phi)\}_{a,a'=L,...,H}$, where:

$$P_i(a'|a;\phi) \propto \exp\left[-\frac{(a'-b_i(a;\phi))^2}{2\sigma_i^2(a;\phi)}\right] \tag{1}$$

is a conditional probability that the i-th group draws ageing outcomes a' given that the true age of the animal is a, and $\sum_{a'=L}^{H} P_i(a'|a;\phi) = 1$ for all a (Punt *et al*. 2008).

The expected age for Reader 1 is assumed to be proportional to the true age:

$$b_1(a) = (1+x)a\,. \tag{2}$$

On the other hand, the expected age for the readers in Group 2 is a linear function of true age a:

$$b_2(a;\phi) = b_L + (b_H - b_L)\frac{a-L}{H-L}\,. \tag{3}$$

This is a 2-parameter model from Punt *et al*. (2008). The parameters in equation (3) should relate to each reader/period when considering hypotheses related to reader/period effects. The values of L and H must be pre-specified and are not estimable parameters (here L =0 and H=70).

The functional form of the ageing error standard deviation for the two groups is also assumed to be a linear function of true age:

$$\sigma_i(a;\phi) = \sigma_{iL} + (\sigma_{iH} - \sigma_{iL})\frac{a-L}{H-L} \qquad (i=1,2)\,. \tag{4}$$

As for the expectation, the parameters in the equation (4) are specific to the reader/period concerned.

Let $\beta = (\beta_L, \ldots, \beta_H)$ be the true age composition of sampled animals, which is unknown. Given the true age (say a), for the *j*-th animal, the contribution of *j*-th sample by Reader 1 to the likelihood is:

$$P_1(a_{1j} \mid a;\phi) = \sum_{k=1}^{r_j} P_1(a_{1jk} \mid a;\phi), \tag{5}$$

where $a_{1j} = (a_{1j1}, \ldots, a_{1jr_j})$. By considering the distribution for Group 2 in a similar way, the joint probability distribution of ageing outcomes by the two groups is provided by a mixture form as

$$\Pr(a_{1j}, a_{2j}; \phi, \beta) = \sum_{a=L}^{H} \beta_a \, P_1(a_{1j} \mid a;\phi) P_2(a_{2j} \mid a;\phi), \tag{6}$$

and therefore the likelihood function for the parameters is given by:

$$Like(\phi,\beta) = \prod_{j=1}^{n} \Pr(a_{1j}, a_{2j}; \phi, \beta) = \prod_{j=1}^{n} \sum_{a=L}^{H} \beta_a \, P_1(a_{1j} \mid a;\phi) P_2(a_{2j} \mid a;\phi). \tag{7}$$

The likelihood contribution for the data that are not in the "valid" category can be expressed as follows; for example, when a data type is "interval" as $[a_{1jk}^{(1)}, a_{1jk}^{(2)}]$, the distribution is:

$$P_1([a_{1jk}^{(1)}, a_{1jk}^{(2)}] \mid a;\phi) = \sum_{a'=a_{1jk}^{(1)}}^{a_{1jk}^{(2)}} P_1(a' \mid a;\phi). \tag{8}$$

In the model above, the parameters in the expectation and variance structures are of interest, while $\beta_L, \ldots, \beta_H$ are nuisance parameters. To make the estimation easier and to reduce the number of nuisance parameters, a functional constraint is incorporated on the parameters for the true age composition of the sample $\beta_a \; (a \ge A)$ as $\beta_a = \beta_A \exp(-Z(a-A))$, where *A* is the largest number which satisfies:

$$\frac{\#\{ j = 1, \ldots, n \mid a_{1j1} \ge A\}}{n} > q, \tag{9}$$

and *Z* is a mortality parameter. The threshold value *q* is, of course, *ad hoc*, but the constraint is nevertheless useful in cases such as this experiment. We use a value *q* = 0.20 as a base case assumption.

Scenarios

We consider a total of four scenarios (Table 3). A main point to be assessed in the sensitivity tests is the bias of the control reader (i.e. Lockyer's bias). For this purpose, we consider the following three cases:

- Case 1: Lockyer is unbiased (*x*=0),
- Case 2: Lockyer provides age-estimates with 10% positively bias (*x*=0.1), and
- Case 3: Lockyer provides age-estimates with 10% negatively biased (*x*=-0.1).

Furthermore, we examine the sensitivity of the results to using all of the data and not just the "valid" estimates (Case 4) for the model with reader-effects. Several alternative models based on the covariates included in the models for the mean and variance structures for age-reading Group 2 are considered (Table 4).

RESULTS AND DISCUSSION

Histograms and scatter plots of the "valid" age-reading outcomes from Lockyer do not suggest evidence for between-trial bias (Figure 1). Similarly, there is no evidence for between-trial bias for Zenitani (Figure 2). Consequently, no covariates for trial are considered in the analyses

Figure 3 plots the age-estimates for each of the Japanese readers ("best" estimates for Masaki and Kato, and the median of the three estimates from Zenitani) against the age-estimates by Lockyer. These plots indicate some discrepancy between the age-estimate obtained by Lockyer and those obtained by the Japanese scientists.

Table 5 summarizes the results of the parameter estimation and model selection for the models with reader effects. Incorporating a reader effect into the variance component (i.e. the extent of random age-reading error) tended to improve the goodness of fit substantially (in terms of model selection criteria) compared to incorporating these effects into the mean structure. Throughout the four cases, Model 3, in which the reader effects were incorporated into both the mean and variance structures, provided the most parsimonious fit to the data. The difference in parameter estimates between Cases 1 (base case) and 4 (use data for index 0-3) was almost negligible, except for Reader 2-2 (Kato). The adequateness of the fits for Model 3 in Case 1 is confirmed by the diagnosis plots shown in Figure 4, where the control reader is assumed to be an unbiased reader.

Table 6 provides the results for models with period (rather than reader) effects. Convergence was not achieved for some models with period effects owing to the large numbers of parameters. Compared to the reader effect models, the period effect models tended not to fit the data as well. This seems to be as expected because the period effects potentially mix reader-effects. There is some evidence for “learning” (in terms of the reduction of variance in ageing) between Periods 4 and 5 for the Japanese third reader (Zenitani).

Overall, the results suggest that the age-reading errors for Lockyer and the three Japanese readers differ. Tables 5 and 6 provide estimates of parameters which could be used to compute ageing error matrices. Ageing-error matrices based on Model 3 (and the assumption of reader rather than period effects) could incorporated into assessments of the impact of age-determination error on the outputs from age-structured models for Antarctic minke whales (e.g. Punt (2010)). It should be noted, however, that although the analyses of this paper are predicated on Lockyer’s age-estimates, it should not necessarily be assumed that Lockyer provides unbiased estimates of age. Rather, the results in tables 5 and 6 provide estimates related to ageing bias and ageing imprecision given different levels of ageing bias by Lockyer.

ACKNOWLEDGEMENTS

The authors thank Christina Lockyer for her excellent work in Japan. The authors are also grateful to Hidehiro Kato, Ryoko Zenitani, Toshiya Kishiro, Hikari Maeda and Yui Zennyoji for their preparation of Lockyer’s age-reading experiment according to the protocol. Hidehiro Kato and Ryoko Zenitai are especially thanked for their dedicated coordination of the experiment and giving the authors helpful information on past Japanese age-reading experiments.

REFERENCES

International Whaling Commission. 2009. Report of the Sub-Committee on In-depth Assessments. Annex G to the Report of the Scientific Committee. *J. Cet. Res. Manage*. 11(Suppl):193-211.

Kitakado, T. 2008. Evaluation of statistical power of a test for examining bias in age-reading experiment for Antarctic minke whales. Paper SC/60/IA16.

Lockyer, C. 2010. Report of the Antarctic minke whale ear plug experiment. Paper SC/62/IA11.

Mori, M., Butterworth, D.S. and Kitakado, T. 2007. Further progress on application of ADAPT-VPA to Antarctic minke whales. Paper SC/59/IA13.

Punt, A.E. 2010. A note on the impact of accounting for age-determination error on the outcome of the statistical catch-at-age analysis for Antarctic minke whales. Paper SC/62/IA6.

Punt, A.E. and Polacheck, T. 2007. Further development of statistical catch-at-age models for southern hemisphere minke whales. Paper SC/59/IA4.

Punt, A.E., Smith, D.C., KrusicGolub, K. and Robertson, S. 2008. Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australia’s Southern and Eastern Scalefish and Shark Fishery. *Can. J. Fish. Aquat. Sci.* 65:1991-2005.

Table 1. The number of samples employed in this experiment

	Period	Group 1	Group 2		
		Lockyer*	Masaki	Kato	Zenitani
Period I	74/75-76/77	50 (10)	50	0	0
Period II	82/83-84/85	50 (10)	0	50	0
Period III	89/90-91/92	50 (10)	0	28	22
Period IV	97/98-99/00	50 (10)	0	0	50
Period V	03/04-05/06	50 (10)	0	0	50

* The numbers in brackets indicate how many plugs were read three times by Lockyer.

Table 2. Types of data given by Lockyer

Index	**Category**[$]	**Data type**	**1st trial***	**2nd trial***	**3rd trial***
0	valid	Age	228 (91.2%)	216 (86.4%)	43 (86%)
1	either	Age1 or Age 2	2 (0.8%)	4 (1.6%)	0
2	minimum	Age >=	11 (4.4%)	10 (4.0%)	3 (6.0%)
3	interval	(Age1, Age2)	1 (0.4%)	0	0
4	may be missing	Age	1 (0.4%)	0	0
10	uncertain	Age	2 (0.8%)	12 (4.8%)	3 (6.0%)
100	unreadable	NA	5 (2.0%)	8 (3.2%)	1 (2.0%)

[$] “valid”: one ageing observation was recorded; “either”: two possible ages were offered; “minimum”: only a minimum age was counted; “interval”: a possible range of ages were given; “missing”: part of plug was missing; “uncertain”: the reader is not confident in the counting.
The numbers in brackets are proportions in percentage in each trial.

Table 3. Scenarios considered in the analyses.

	Covariate	**Bias in control reader**	**Data**
Case 1 (Base)	Reader or Period effects	0%	"valid" only
Case 2	Reader or Period effects	10%	"valid" only
Case 3	Reader or Period effects	-10%	"valid" only
Case 4	Reader effects	0%	Index=0,1,2,3

Table 4. Assumption regarding the covariate effects and parameters

Model	Description
0	No reader /period effects
1	Reader /period effects in the mean structure
2	Reader /period effects in the variance structure (constrained by $\sigma_{iL} \leq \sigma_{iH}$)
2c	Reader/period effects in the variance structure (σ_{2L} is common among readers in Group 2)*
3	Reader /period effects in the mean and variance structures (constrained by $\sigma_{iL} \leq \sigma_{iH}$)
3c	Reader/period effects in the mean and variance structures (σ_{2L} is common among readers in Group 2)*

*To reduce the number of parameters

Table 5. Results for the reader effects (Upper: estimate, Lower: SE). The column "#parameters" indicates the number of non-nuisance parameters.

Case 1. Lockyer's percent bias = 0, L=0, H=70, q =0.2

Model	Loglike	#parameters	Δ-AIC	Δ-AICc	Reader 1 (Lockyer)		Reader 2-1 (Masaki)		Reader 2-2 (Kato)		Reader 2-3 (Zenitani)		Reader 2-1 (Masaki)		Reader 2-2 (Kato)		Reader 2-3 (Zenitani)	
					sigL1	sigH1	bL21	bH21	bL22	bH22	bL23	bH23	sigL21	sigH21	sigL22	sigH22	sigL23	sigH23
0	-2441.13	6	68.52	51.19	1.67	2.92	1.41	61.26					0.35	4.78				
					0.17	0.84	0.10	0.50					0.07	0.65				
1	-2425.79	10	45.84	37.14	1.57	3.89	2.51	60.98	2.06	56.72	1.23	62.13	0.43	3.76				
					0.18	0.80	0.44	2.00	0.38	1.25	0.11	0.67	0.06	0.48				
2	-2409.60	10	13.46	4.76	1.51	3.60	1.28	61.83					1.82	1.82	0.16	8.85	0.46	2.87
					0.16	0.55	0.11	0.53					0.26	0.26	0.38	1.53	0.06	0.33
2c	-2415.87	8	22.00	8.97	1.55	3.64	1.30	61.92					0.46	6.29		7.98		2.82
					0.17	0.57	0.11	0.55					0.06	1.76		0.97		0.32
3	-2398.87	14	0.00	0.00	1.42	3.75	3.16	58.50	2.64	55.47	1.19	62.03	1.58	1.59	0.82	6.60	0.47	2.85
					0.16	0.55	0.53	1.94	0.58	1.94	0.10	0.57	0.22	0.22	0.39	1.30	0.06	0.32
3c	-2404.39	12	7.04	2.68	1.48	3.69	2.73	60.17	2.31	56.40	1.19	62.06	0.49	5.80		7.39		2.77
					0.16	0.55	0.60	2.77	0.48	1.82	0.10	0.58	0.06	1.46		0.93		0.33

Case 2. Lockyer's percent bias = 0.1, L=0, H=70, q =0.2

Model	Loglike	#parameters	Δ-AIC	Δ-AICc	Reader 1 (Lockyer)		Reader 2-1 (Masaki)		Reader 2-2 (Kato)		Reader 2-3 (Zenitani)		Reader 2-1 (Masaki)		Reader 2-2 (Kato)		Reader 2-3 (Zenitani)	
					sigL1	sigH1	bL21	bH21	bL22	bH22	bL23	bH23	sigL21	sigH21	sigL22	sigH22	sigL23	sigH23
0	-2437.39	6	64.60	47.27	1.55	3.36	1.56	66.93					0.38	5.15				
					0.17	0.96	0.09	0.57					0.09	0.72				
1	-2423.94	10	45.70	37.00	1.46	4.44	2.39	66.96	2.08	62.07	1.33	68.31	0.46	4.04				
					0.19	1.01	0.47	2.31	0.36	1.37	0.11	0.75	0.08	0.62				
2	-2409.60	10	17.02	8.32	1.38	4.26	1.53	67.40					1.74	1.75	0.19	9.60	0.52	2.90
					0.15	0.64	0.13	0.61					0.26	0.26	0.42	1.83	0.07	0.39
2c	-2412.61	8	19.04	6.01	1.42	4.31	1.55	67.42					0.53	6.05		8.50		2.82
					0.16	0.63	0.12	0.61					0.07	1.98		1.08		0.36
3	-2397.09	14	0.00	0.00	1.32	4.33	3.14	64.10	2.70	60.59	1.33	68.11	1.57	1.57	0.82	7.08	0.51	2.95
					0.17	0.67	0.53	2.21	0.54	2.07	0.41	1.16	0.23	0.23	0.36	1.44	0.07	0.40
3c	-2402.49	12	6.80	2.44	1.35	4.37	2.98	64.73	2.47	61.31	1.37	68.01	0.56	6.32		7.83		2.74
					0.17	0.64	0.71	3.47	0.53	2.15	0.16	0.71	0.08	1.74		1.04		0.36

Table 5 (continued)

Case 3. Lockyer's percent bias = - 0.1, L=0, H=70, q =0.2

Model	Loglike	#parameters	Δ-AIC	Δ-AICc	Reader 1 (Lockyer)		Reader 2-1 (Masaki)		Reader 2-2 (Kato)		Reader 2-3 (Zenitani)		Reader 2-1 (Masaki)		Reader 2-2 (Kato)		Reader 2-3 (Zenitani)	
					sigL1	sigH1	bL21	bH21	bL22	bH22	bL23	bH23	sigL21	sigH21	sigL22	sigH22	sigL23	sigH23
0	-2436.43	6	63.96	46.63	1.53	3.20	1.55	55.15					0.40	4.13				
					0.17	0.81	0.10	0.49					0.08	0.57				
1	-2422.53	10	44.16	35.46	1.49	3.68	2.44	55.23	2.02	51.28	1.44	55.93	0.44	3.53				
					0.19	0.85	0.46	1.83	0.37	1.10	0.12	0.55	0.08	0.56				
2	-2406.15	10	11.40	2.70	1.40	3.57	1.49	55.46					1.77	1.78	0.18	7.91	0.48	2.60
					0.15	0.49	0.12	0.47					0.25	0.25	0.48	1.50	0.08	0.32
2c	-2411.77	8	18.64	5.61	1.41	3.66	1.54	55.44					0.52	5.49		7.11		2.51
					0.16	0.50	0.17	0.52					0.10	1.51		0.88		0.33
3	-2396.45	14	0.00	0.00	1.36	3.65	3.16	52.97	2.68	50.10	1.40	55.78	1.58	1.59	0.83	5.93	0.48	2.60
					0.15	0.48	0.53	1.72	0.59	1.71	0.11	0.48	0.22	0.22	0.40	1.16	0.07	0.30
3c	-2401.64	12	6.38	2.02	1.41	3.62	2.76	54.32	2.39	50.80	1.40	55.79	0.51	5.29		6.59		2.51
					0.15	0.49	0.57	2.39	0.47	1.57	0.11	0.49	0.08	1.31		0.84		0.30

Case 4. Lockyer's percent bias = 0, L=0, H=70, q =0.2 (using data with Index=0,1,2,3)

Model	Loglike	#parameters	Δ-AIC	Δ-AICc	Reader 1 (Lockyer)		Reader 2-1 (Masaki)		Reader 2-2 (Kato)		Reader 2-3 (Zenitani)		Reader 2-1 (Masaki)		Reader 2-2 (Kato)		Reader 2-3 (Zenitani)	
					sigL1	sigH1	bL21	bH21	bL22	bH22	bL23	bH23	sigL21	sigH21	sigL22	sigH22	sigL23	sigH23
0	-2464.58	6	65.02	47.73	1.69	3.31	1.37	61.14					0.37	4.37				
					0.21	1.72	0.13	0.52					0.09	1.21				
1	-2448.33	10	40.52	31.84	1.63	3.74	2.51	60.91	2.01	56.60	1.15	62.17	0.42	3.76				
					0.18	0.75	0.47	2.04	0.39	1.25	0.15	0.68	0.06	0.45				
2	-2435.42	10	14.70	6.02	1.57	3.50	1.27	61.75					1.76	1.76	0.14	8.73	0.45	2.87
					0.16	0.54	0.10	0.52					0.26	0.26	0.37	1.50	0.06	0.33
2c	-2440.91	8	21.68	8.68	1.61	3.55	1.29	61.83					0.46	5.94		7.82		2.82
					0.16	0.56	0.11	0.53					0.06	1.69		0.95		0.32
3	-2424.07	14	0.00	0.00	1.52	3.56	3.07	58.78	2.23	58.78	1.11	62.15	1.53	1.53	0.47	7.24	0.45	2.90
					0.16	0.54	0.52	1.91	0.53	1.91	0.12	0.57	0.22	0.22	0.45	1.57	0.06	0.33
3c	-2428.58	12	5.02	0.67	1.55	3.58	2.67	60.35	2.27	56.21	1.12	62.14	0.47	5.65		7.22		2.82
					0.16	0.55	0.52	2.48	0.46	1.75	0.13	0.58	0.06	1.42		0.92		0.33

Table 6. Results for the period effects models. Convergence was not reached for some models owing to the large number of parameters (“Hessian”: the hessian was degenerated; “NC”: not convergence). Δ-AIC and Δ-AICc are the difference from the best model in each case shown Table 5.

Case 1. Lockyer's percent bias = 0, L=0, H=70, q =0.2

Model	Loglike	#parameters	Δ-AIC	Δ-AICc	Reader 1 (Lockyer)		Period 1		Period 2		Period 3		Period 4		Period 5		Period 1		Period 2		Period 3		Period 4		Period 5	
					sigL1	sigH1	bL21	bH21	bL22	bH22	bL23	bH23	bL24	bH24	bL25	bH25	sigL21	sigH21	sigL22	sigH22	sigL23	sigH23	sigL24	sigH24	sigL25	sigH25
0	-2441.1	6	68.52	51.19	1.67	2.92	1.41	61.26									0.35	4.78								
					0.17	0.84	0.10	0.50									0.07	0.65								
1	-2416.4	14	35.06	35.06	1.57	3.69	2.35	61.51	2.13	58.01	1.62	57.19	1.16	62.49	1.86	62.28	0.39	3.99								
					0.18	0.71	0.43	2.05	0.55	1.73	0.32	1.20	0.12	1.07	0.15	0.88	0.05	0.42								
2	-2413.1	14	28.44	28.44	1.36	3.28	1.86	61.11									1.68	1.70	0.00	7.85	0.06	8.70	0.64	3.66	0.44	2.67
	(Hessian)																									
2c	-2422.6	10	39.50	30.80	1.48	3.30	1.28	61.76									0.44	6.76		6.85		6.69		3.94		2.55
					0.14	0.53	0.13	0.52									0.06	1.56		1.08		0.92		0.52		0.38
3	NA	22																								
	(NC)																									
3c	-2403.5	18	17.30	26.08	1.46	3.32	2.64	60.50	2.36	57.42	1.84	56.62	1.18	62.66	1.83	62.15	0.42	6.05		6.56		6.04		3.99		2.61
					0.16	0.53	0.72	3.12	0.72	2.28	0.36	1.44	0.14	1.13	0.15	0.74	0.06	1.66		1.06		0.84		0.55		0.38

Case 2. Lockyer's percent bias = 0.1, L=0, H=70, q =0.2

Model	Loglike	#parameters	Δ-AIC	Δ-AICc	Reader 1 (Lockyer)		Period 1		Period 2		Period 3		Period 4		Period 5		Period 1		Period 2		Period 3		Period 4		Period 5	
					sigL1	sigH1	bL21	bH21	bL22	bH22	bL23	bH23	bL24	bH24	bL25	bH25	sigL21	sigH21	sigL22	sigH22	sigL23	sigH23	sigL24	sigH24	sigL25	sigH25
0	-2437.39	6	64.60	47.27	1.55	3.36	1.56	66.93									0.38	5.15								
					0.17	0.96	0.09	0.57									0.09	0.72								
1	-2419.68	14	45.18	45.18	1.72	3.54	2.59	66.72	2.39	63.01	1.70	62.72	1.12	68.45	2.55	66.80	0.39	4.35								
					0.19	0.82	0.36	2.06	0.61	2.03	0.31	1.31	0.14	1.11	0.13	0.90	0.05	0.50								
2	NA	14																								
	(NC)																									
2c	-2418.11	10	34.04	25.34	1.37	3.80	1.56	67.30									0.47	6.73		7.24		7.56		4.38		2.54
					0.13	0.59	0.10	0.58									0.07	1.75		1.21		1.07		0.63		0.42
3	NA	22																								
	(NC)																									
3c	-2401.8	18	17.42	26.20	1.34	3.80	3.13	64.04	2.24	63.22	1.76	62.11	1.18	69.47	1.59	67.80	0.46	7.43		6.95		6.54		4.35		2.60
					0.14	0.59	0.72	3.69	0.76	2.64	0.38	1.69	0.15	1.12	0.16	0.83	0.06	1.85		1.19		0.94		0.63		0.43

Case 3. Lockyer's percent bias = - 0.1, L=0, H=70, q =0.2

Model	Loglike	#parameters	Δ-AIC	Δ-AICc	Reader 1 (Lockyer)		Period 1		Period 2		Period 3		Period 4		Period 5		Period 1		Period 2		Period 3		Period 4		Period 5	
					sigL1	sigH1	bL21	bH21	bL22	bH22	bL23	bH23	bL24	bH24	bL25	bH25	sigL21	sigH21	sigL22	sigH22	sigL23	sigH23	sigL24	sigH24	sigL25	sigH25
0	-2436.43	6	63.96	46.63	1.53	3.20	1.55	55.15									0.40	4.13								
					0.17	0.81	0.10	0.49									0.08	0.57								
1	-2416.58	14	40.26	40.26	1.51	3.22	2.44	55.23	2.08	52.35	1.60	51.72	1.35	56.71	1.61	56.25	0.40	3.91								
					0.17	0.71	0.45	1.82	0.60	1.58	0.42	1.12	0.17	1.14	0.15	0.88	0.07	0.55								
2	-2409.41	14	25.92	25.92	1.36	3.21	1.54	55.39									1.76	1.79	0.01	6.86	0.15	7.17	0.48	3.59	0.42	2.32
	(Hessian)																									
2c	-2416.65	10	32.40	23.70	1.38	3.29	1.54	55.46									0.45	5.85		6.08		6.25		3.72		2.23
					0.14	0.47	0.11	0.63									0.07	1.41		0.98		0.90		0.52		0.33
3	NA	22																								
	(NC)																									
3c	-2401.06	18	17.22	26.00	1.37	3.26	2.68	54.61	2.28	52.03	1.90	51.02	1.33	56.59	1.55	56.06	0.44	5.71		5.80		5.59		3.69		2.26
					0.14	0.45	0.55	2.35	0.74	2.00	0.43	1.40	0.13	0.94	0.13	0.67	0.06	1.30		0.94		0.79		0.50		0.32

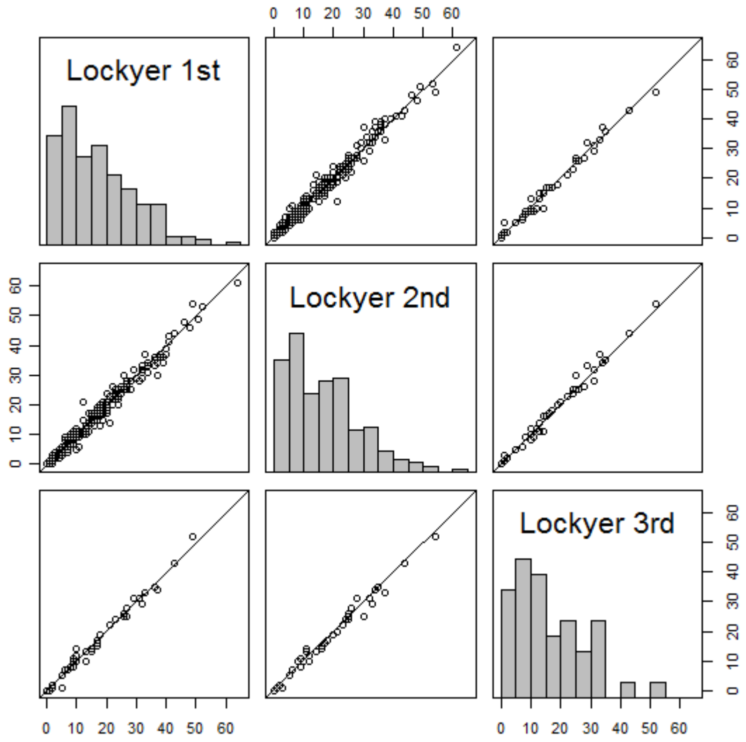

Figure 1. Scatter plots and histograms for Lockyer's ageing data ("valid" data only) for the three trials.

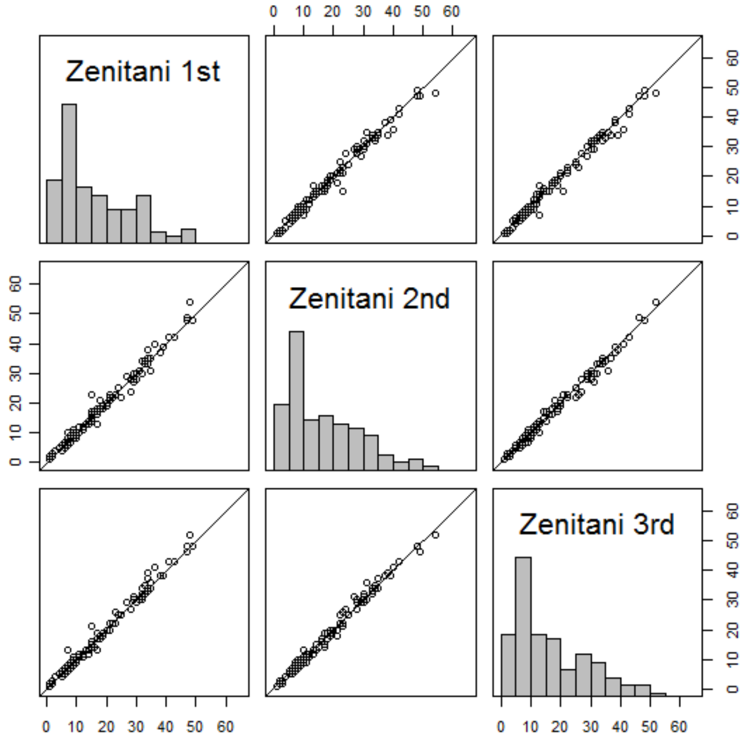

Figure 2. Scatter plots and histograms for Zenitani's ageing data for her three trials.

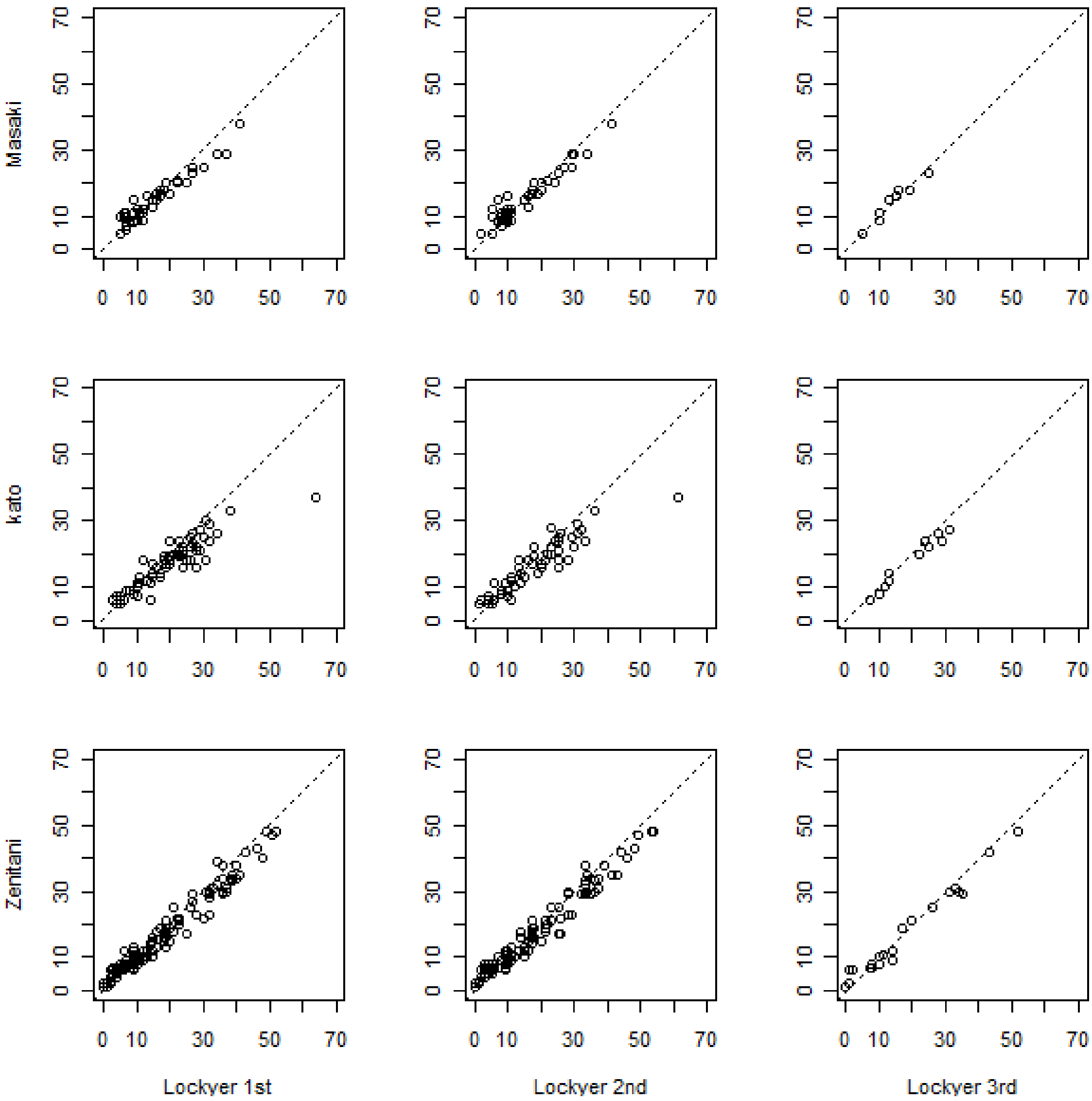

Figure 3. Scatter plots of the “best” age-estimates form the Japanese readers against Lockyer’s 1st, 2nd and 3rd trials (“valid” data only). The dashed lines show a 45 degree line.

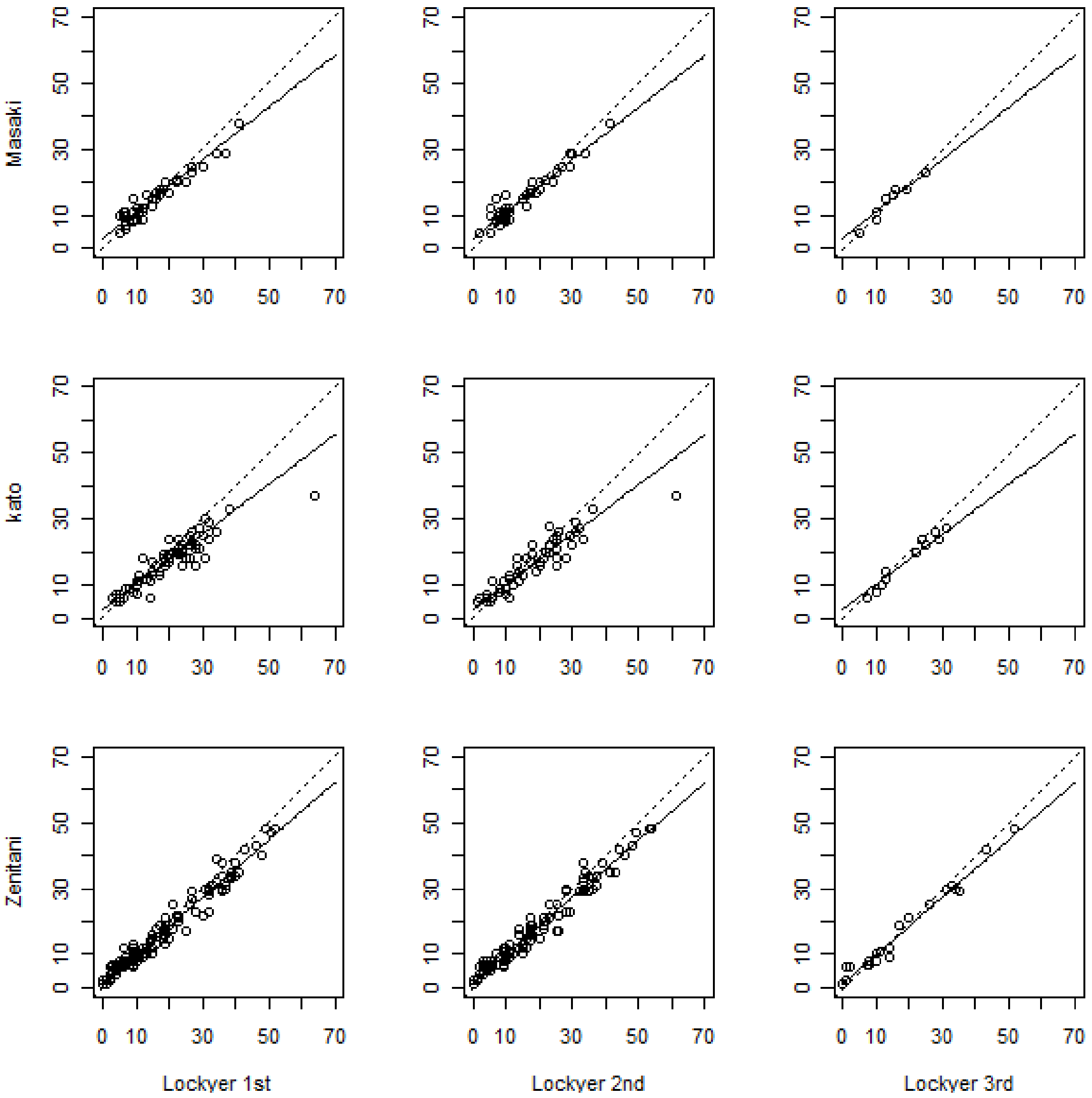

Fig 4. Diagnosis plots for the best model (Model 3) in Case 1 (the control is an unbiased reader). Solid lines show the estimates of expected ages given in equation (3) for the Japanese readers.