# Appendix S1

Calculations for Perinate Age in Foetal Weeks

The high number of perinates at Khok Phanom Di lent itself to a closer examination of age, (Franklin, 2010; Cunningham et al., 2016). In addition to Tayles’s (1999) age estimates, a further assessment of age was made for perinatal individuals by calculating foetal weeks from diaphyseal measurements recorded from Halcrow et al. (2008) and Tayles (1999). It is difficult to calculate an accurate length of gestation for new-borns; as Cunningham et al. point out, separate disciplines record age differently, and the weight and length of a full-term baby are population dependent (2016, 7-8).

Age in foetal weeks was calculated for perinates from diaphyseal long bone measurements reported in both Halcrow et al., (2008) and Tayles (1999). The latter published maximum diaphyseal lengths of infants, children and adolescents. Halcrow and colleagues (2008) published the diaphyseal lengths for 49 perinatal individuals from Khok Phanom Di, comparing age distribution to other Thai sites[[1]](#footnote-1). Halcrow et al. (2008) took measurements in millimetres, to one decimal place, using the Buikstra and Ubelaker (1994) standard with Mitutoyo digital callipers. Tayles (1999) also recorded to one decimal place[[2]](#footnote-2). The individuals measured by Tayles (1999) and Halcrow et al. (2008) largely overlapped; however, not entirely. In order to age more perinate individuals, it was desirable to combine the datasets. However, if there had been significant interobserver variability between the two sets of measurements it would increase error for individuals where only one set of measurements existed. In order to establish whether the two datasets could be combined it was necessary to confirm if there was a statistically significant difference or a difference in equivalence between the measurements taken by both authors. The measurements suitable for comparison were the diaphyseal length of the humerus, ulna, radius, femur and tibia. It was not possible to access whether measurements of the fibulae were significantly different as only Halcrow et al. (2008) had included them in their study. Comparison of 44 individuals was carried out using a paired samples t-test for the humerus, radius, and tibia as they met the assumption of normal distribution; and a Wilcoxon signed rank test for the ulna, and femur as they did not have a normal distribution; with significance set at 95% (p value <0.05). The paired sample t-tests of the humeri suggests a significant difference between the measurements made by the two authors different (t=2.59, df=32, *p*=0.01). However, the tests suggest there was not a significant difference between radial (t=-1.08, df=22, *p*=0.29), femoral (t=-0.86, df=33, *p*=0.40), and tibial (t=1.81, df=32, *p*=0.08) measurements taken by Tayles (1999) and Halcrow et al. (2008). The Wilcoxon signed rank test suggests that there is no significant difference between the ulna (V=110, *p*=0.56) and the femoral (V=142.5, *p*=0.41) measurements in the two publications. A paired two one-sided test (TOST) for equivalence was also carried out on the five pairs of measurements, with a 95% TOST interval and tolerance of 1mm[[3]](#footnote-3) (humerus *p*=0.40; radius *p*=2.25e-07; ulna *p*=0.00; femur *p*=3.44e-16; tibia *p*=0.00). The null hypothesis here is of statistical difference, which was rejected in all cases. This would indicate that the measurements from Tayles (1999) and Halcrow et al., (2008) for the radius, ulna, femur and tibia were not significantly different and were statistically equivalent; the humeral measurements; however, were significantly different but were statistically equivalent. The statistical equivalence of all bone measurements, indicating a minimal degree of interobserver error, was felt to be sufficient to combine them, taking an average, excluding any outliers in difference[[4]](#footnote-4), and include the Halcrow et al., (2008) fibula measurements in order to calculate foetal weeks.

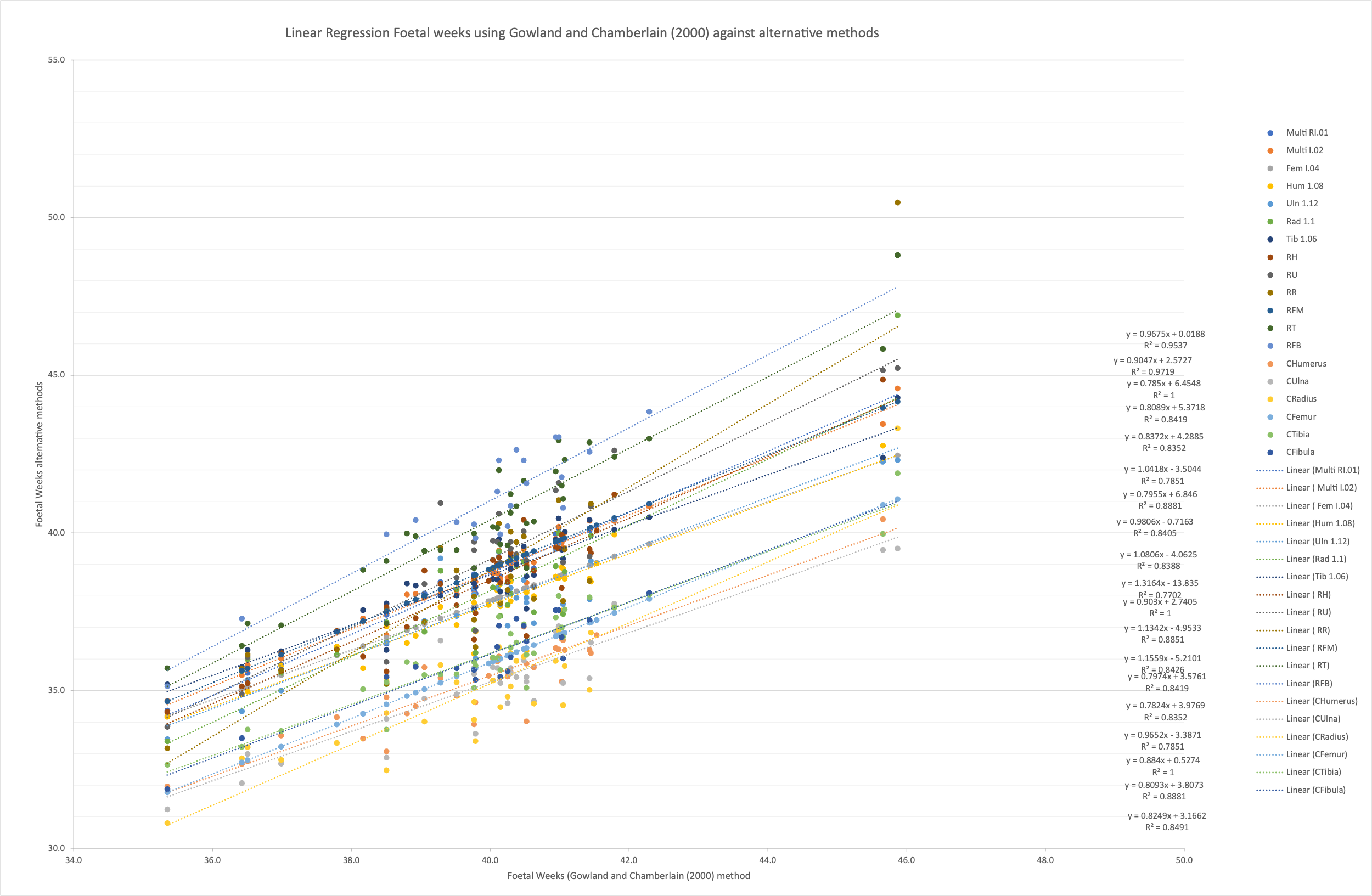
Using the new combined long bone measurements, it was possible to estimate gestational age for 53 individuals[[5]](#footnote-5), using the multi-method approach (similar to Halcrow et al. 2008). This combined means from seven calculations from Scheuer et al. (1980) and Scheuer and Black, (2000)[[6]](#footnote-6); six methods from Sherwood et al. (2000)[[7]](#footnote-7); and six methods from Carneiro et al. (2016)[[8]](#footnote-8). These authors used foetuses of known age to develop regression models based on different long bones. The femur is identified as the bone providing the highest level of accuracy; however, a multi-method approach enables calculations to be made for incomplete skeletons, where for example the femora may not be present. In addition to these regression-based methods, a further assessment of age was made taking the combined measurements for the femora and applying a Bayesian procedure developed by Gowland and Chamberlain (2002)[[9]](#footnote-9). Chamberlain (2000) in earlier work argued that a Bayesian estimation method can remove population bias. Gowland and Chamberlain (2002) argue that previous regression models generate an artificially narrow age distribution because of the non-uniformity of the age distribution from the reference samples used to develop regression-based methods. The authors counter this by accounting for the different probability of death at each week of gestation by incorporating the prior probability of mortality risk at a given age-at-death based on foetal and infant diaphyseal length data from clinical studies of individuals with known gestational age (studies used: Maresh and Deming, 1939; Maresh, 1955; Fazekas and Kósa, 1978; Scheuer et al., 1980; Jeanty et al., 1981; 1982). However, while providing a more reliable representation of the age distribution, Gowland and Chamberlain’s method is limited by its use of only one bony element.

In literature reporting age in foetal weeks, authors will commonly report the foetal week estimations from the different methods and take an average in order to gain a single estimate for perinate age (Halcrow et al., 2008; Kinaston, et al., 2009). The advantage of using all nineteen methods is that it maximises data collection for incomplete individuals; however, this is problematic with a data set that includes missing values, in much the same way as combining Tayles’s (1999) and Halcrow et al.’s (2008) measurements discussed above. In order to compare the twenty different methods across the three different authors an analysis of variance (ANOVA) was performed on a linear mixed effects model (lmer) of the twenty methods. This determined a highly significant difference between them (p=2.2e-16). An estimated marginal means (EMMs) test was then performed, which demonstrated the relationship between each method. Of the 190 pairs, only 34 were not significantly different.[[10]](#footnote-10) Similarly, the same methods applied to the different methods grouped by author resulted in a significant difference between the authors (all six pairs p==<0.0001). The difference between the methods leaves three options to achieve a single estimate for age in foetal weeks. The first is to take an average, acknowledging a bias based on the number and particular methods used to make the estimate, this would make it difficult to achieve an accurate relative comparison of the perinates at Khok Phanom Di. A second option would be to choose a preferred method for estimation; however, this would result in a loss of individuals who do not have the bony element(s) required for the calculation (Barnard and Meng, 1999; Horton and Kleinman, 2007). Alternatively, choose a preferred method and impute the data from the other methods using a regression equation. This last option would create less variability in the data and provide an estimate of age suitable for a relative comparison.

The CIfA suggest a policy of inclusion for skeletal material used in analysis, as they state that exclusion of skeletons with less-than-ideal preservation “may lead to the loss of significant information of pathological conditions that result in the loss of bone density and structure” (Brickley, 2017, 7-8). For the purposes of this study ascertaining age in foetal weeks would only be used for relative comparison within the Khok Phanom Di assemblage. As such it was important to find a balance between the quantity of data and its relative accuracy. Imputation was felt to be the most inclusive method for calculating perinate age.

Imputation of the data was done by selecting Gowland and Chamberlain (2002) as the preferred method for estimating foetal weeks. This method was selected because of the prior probability inclusion described above. A linear regression was calculated for each of the other nineteen methods against Gowland and Chamberlain, excluding individuals over 50 weeks, following Halcrow et al. (2008). The equation for the regression model from each was used to impute the ages from the other nineteen methods to achieve an estimate comparable to Gowland and Chamberlain for the missing values (Figure S1). A final estimate for age in foetal weeks was calculated from Gowland and Chamberlain with missing values imputed from an average of the available Scheuer and Black (2000) and Sherwood et al., (2000) calculations using the combined measurements from both Tayles (1999) and Halcrow et al. (2008). This enabled foetal week calculations for 46 of the 49 perinates available for analysis, measurements were unavailable for the remaining three individuals. The difference in distribution of estimated age between foetal weeks calculated the mean of all twenty methods versus the imputed values shows a broader and younger distribution in age determined by the collective methods (Figure S2).

It should be noted that the methods use modern and historical European data sets to develop their calculations, this may not be applicable to prehistoric Southeast Asian groups. Gowland and Chamberlain (2000, 681-682) note this, but also cite Allen et al. (2000), who found that different socio-economic groups in the USA had different levels of mortality; however, they did have a similar age-mortality distribution. A difference in populations would be more problematic if comparing between different cultural groups; however, the use here provides a relative comparison between the perinates at Khok Phanom Di and the relationship to ochre. As noted previously, Cox (2000) describes an ‘obsession’ with age and the cultural bias and assertions that go with it. The efforts to assign age made here provide a ‘rank’ order of perinates by age, the assignment of a number of foetal weeks is made with caution. Any interpretation made based on age in foetal weeks is considering whether individuals are older or younger than others within their society and not compared to other groups.

*Figure S1* Linear regression of imputed foetal week values using Gowland and Chamberlain (2000) against other methods.

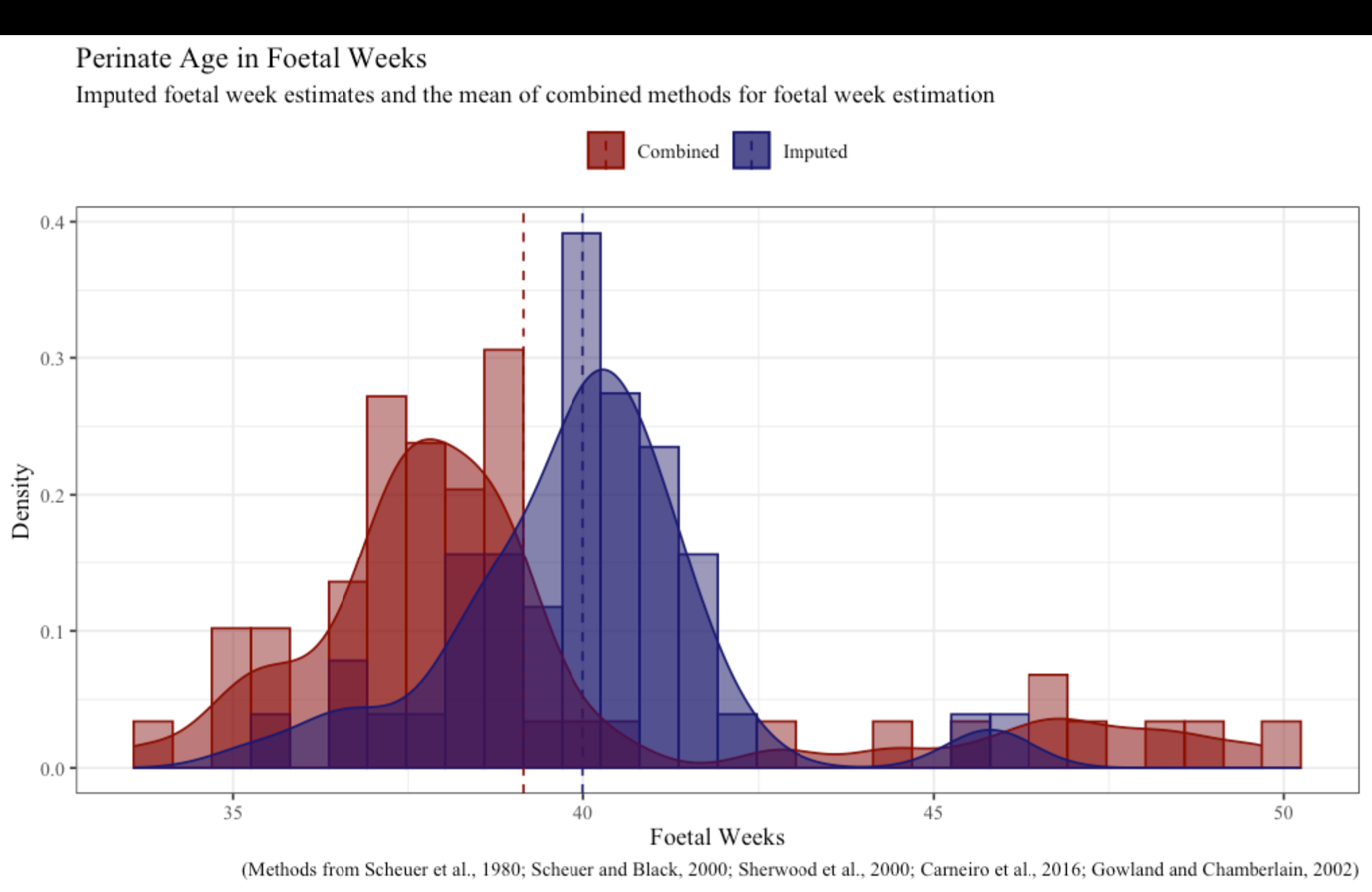


Figure S2. Histogram showing the imputed and combined methods for foetal week values. The dashed lines show the mean is lower for the combined than the imputed age in foetal weeks.

Gowland and Chamberlain (2000)

GCF = (0.4208(F)) +9.0467 ± (see graph below)

Scheuer et al., 1980; Scheuer and Black 2000 (Sc)

Sc1.01 = (0.1724(F)+0.1538(T)+0.0674(H)-0.0718(R)+0.1397(U)+7.2624) ±1.88

Sc1.02 = (0.1984(FM)+0.2291(T)+9.3575) ±1.87; Sc1.04 = (0.3303(FM)+13.5583) ±2.08

Sc1.06 = (0.4207(T)+11.4724) ±2.12

Sc1.08 = (0.4585(H)+8.6563) ±2.33

Sc1.10 = (0.5850(R)+7.7100) ±2.29

Sc1.12 = (0.5072(U)+7.8208) ±2.20

Sharwood et al., 2000 (Sh)

ShH = (12.98+0.25(H)+(0.0024)(H2)) ±2.12

ShU = (14.28 +0.19(U)+0.0039)(U2)) ±2.08

ShR = (13.53+0.25(R)+(0.0045)(R2)) ±2.14

ShFM = (10.91+0.38(FM)) ±2.05

ShT = (15.13+0.19(T)+(0.0031)(T2)) ±2.06

ShFB = (14.72+0.21(FB)+0.0036(FB2)) ±2.19

Carneiro et al., 2016 (C)

CH = (0.452(H)+6.814) ±2.31

CU = (0.474(U)+7.278) ±2.23

CR = (0.542(R)+7.003) ±2.48

CF = (0.372(F)+8.525) ±2.00

CT = (0.428(T)+8.514) ±2.12

CFB = (0.451(FB)+8.603) ±2.22.

# F = femoral length; T = tibial length; H = humeral length; R = radial length; U = ulna length; FB fibula length. All long bone measurements in mm and standard error in weeks. Each method name is taken from the original text with the addition of the first author initial.

1. In this work the authors calculated age in foetal weeks, using regression methods 1.01, 1.02, 1.04, 4.01, 4.03 from Scheuer et al., (1980) and Scheuer and Black (2000); and Humerus, Ulna, Radius, Femur, Tibia, Fibula from Sherwood et al., (2000). [↑](#footnote-ref-1)
2. However, only two individuals appear to have measurements that are between half a decimal place, the vast majority of measurements are whole numbers (Tayles, 1999, 54-56). [↑](#footnote-ref-2)
3. This was deemed appropriate as it appears as though the majority of Tayles’s (1999) measurements were taken to the nearest 1mm. [↑](#footnote-ref-3)
4. Outliers difference between the two sets of measurements were considered individually. The radial measurement from B52, B66, and the tibial measurement for B124 from Tayles (1999) were excluded as scale photographs supported the Halcrow et al. (2008). The humeral measurement for B99 from Halcrow et al. (2008) was excluded as scale photographs supported Tayles (1999), likely a misprint as the two sets of measurements were inverted 86mm and 68mm. The radial measurements from both authors were excluded from B66 as they were incomplete on examination and so could not corroborate either measurement. These were removed from all subsequent calculations. [↑](#footnote-ref-4)
5. This number is higher than the total number of perinates from the assemblage. In calculating foetal weeks individuals who were aged below 50 weeks by any of the twenty methods were included. Seven of these individuals were not considered perinatal overall (skeletal and dental morphology) and were aged 0.3 years (2; 82; 89; 97; 106; 133; 133; 134). All seven were excluded in the final estimation of foetal weeks (explanation of the calculations explained in text), but appear graphically in Figure S1. [↑](#footnote-ref-5)
6. Selected methods from Scheuer et al., 1980; Scheuer and Black 2000 (Sc) Sc1.01 = (0.1724(F)+0.1538(T)+0.0674(H)-0.0718(R)+0.1397(U)+7.2624) ±1.88; Sc1.02 = (0.1984(FM)+0.2291(T)+9.3575) ±1.87; Sc1.04 = (0.3303(FM)+13.5583) ±2.08; Sc1.06 = (0.4207(T)+11.4724) ±2.12; Sc1.08 = (0.4585(H)+8.6563) ±2.33; Sc1.10 = (0.5850(R)+7.7100) ±2.29; Sc1.12 = (0.5072(U)+7.8208) ±2.20. F = femoral length; T = tibial length; H = humeral length; R = radial length; U = ulna length. All long bone measurements in mm and standard error in weeks. Each method name is taken from the original text with the addition of the first author initial. [↑](#footnote-ref-6)
7. Methods from Sharwood et al., 2000 (Sh) ShH = (12.98+0.25(H)+(0.0024)(H2)) ±2.12; ShU = (14.28 +0.19(U)+0.0039)(U2)) ±2.08; ShR = (13.53+0.25(R)+(0.0045)(R2)) ±2.14; ShFM = (10.91+0.38(FM)) ±2.05; ShT = (15.13+0.19(T)+(0.0031)(T2)) ±2.06; ShFB = (14.72+0.21(FB)+0.0036(FB2)) ±2.19. Abbreviations as above plus FB = fibula length, long bone measurements in mm and standard error in weeks. Each method name is taken from the long bone used with addition of the first author initial. [↑](#footnote-ref-7)
8. Method from Carneiro et al., 2016 CH = (0.452(H)+6.814) ±2.31; CU = (0.474(U)+7.278) ±2.23; CR = (0.542(R)+7.003) ±2.48; CF = (0.372(F)+8.525) ±2.00; CT = (0.428(T)+8.514) ±2.12; CFB = (0.451(FB)+8.603) ±2.22. Abbreviations as above, long bone measurements in mm and standard error in weeks. Each method name is taken from the long bone used with addition of the first author initial. [↑](#footnote-ref-8)
9. Weighted means from the prior probabilities of age model Gowland and Chamberlain (2002) = (0.4208\*(F))+9.0467). ± values vary depending on the length of the bone, the smallest being ±0 at 15mm and the largest being +5.16/-4.84 at 65mm. Abbreviations as above, long bone measurements in mm and standard error in weeks. [↑](#footnote-ref-9)
10. Gowland:ShT; CF:CFB,CT; CFB:CT; CH:CR,CU; CR:CU; Sc1.05:Sc1.08,Sc1.12,Sc1.10; Sc1.08:Sc1.12, Sc1.10; Sc1.01:Sc1.02, Sc1.06, ShF, ShH, ShR, ShU; Sc1.02:Sc1.06, ShF, ShH, ShR, ShU; Sc1.10:Sc1.12, ShR, ShH; Sc1.06: ShF, ShH, ShR; ShF: ShH, ShR, ShRU; ShH:ShR. [↑](#footnote-ref-10)