Using a probabilistic approach, two morphometric methods are integrated to quantify morphological variability among individuals. The results of these comparisons are expressed as standardised probabilistic Z-scores such that these statistics can serve to quantify variation in the shape of a cranium or mandible in the context of a constellation of landmarks. In this study, PD and log sem statistics are integrated and standardised using Z-scores, and applied probabilistically to Plio-Pleistocene hominins. As a test case, OH 7 and OH 24 as reference specimens of *Homo habilis* are compared to fossils representing other taxa. There is a wide spectrum of variation in Z-scores for specimens attributed to early *Homo* dated within the period between circa 1.8 Ma and 2 Ma. In terms of morphometric variation predating 1.8 Ma, Z-scores (Z<2) for *Australopithecus afarensis*, *A. africanus* and *Homo habilis* display a small range of variability. This study serves as a demonstration of a method whereby log sem and PD can be used together to facilitate an objective assessment of morphological variability, applicable in palaeoanthropological contexts.

**Significance:**

- Using a probabilistic approach, two morphometric methods are integrated to quantify morphological variability in Plio-Pleistocene African hominin mandibles and crania.
- Two Tanzanian specimens of *Homo habilis* (the OH 7 mandible of the holotype specimen, and the OH 24 skull) can be used as reference material for morphometric comparisons with other fossils (mandibles or crania) attributed to *Australopithecus africanus*, *A. afarensis*, *H. erectus* and *H. rudolfensis*.
- The results of these comparisons are expressed as standardised probabilistic Z-scores such that these statistics for skulls and mandibles can be expressed on a common scale.

**Materials**

A hominin mandible catalogued as OH 7 (virtual reconstruction shown in Figure 1), attributed to *H. habilis* (the holotype), was discovered at Olduvai Gorge (Bed I) in Tanzania from deposits dated at 1.8 Million years ago (Ma). OH 24 is a contemporaneous skull of the same species from the same site. The two specimens can be used as reference material for morphometric comparisons with other fossils (mandibles or crania) attributed to *Australopithecus africanus*, *A. afarensis*, *H. erectus* and *H. rudolfensis*.

The materials used in this study (Table 1) relate primarily to fossils from three general time periods: firstly, specimens attributed to *A. afarensis*, circa 3 million years ago; secondly, specimens attributed to *A. africanus*, circa 2.5 Ma; and, thirdly, specimens attributed to early *Homo*, between circa 1.8 Ma and 2 Ma. *A. afarensis* mandibles from Ethiopia0,14-15 include AL 288-1 ("Lucy"), AL 200, AL 822 and AL 400-1. A mandible of *A. africanus* from Sterkfontein in South Africa is Sts 52b, while Sts 5 (the skull of *Mrs Ples*) from the same site also represents the latter species.19 Sts 71 (a partial skull) from Sterkfontein was originally attributed to *A. africanus* but has also been referred to as *A. prometheus*.16-18 AL 668-1 is a maxilla (*Homo sp.*) from Hadar in Ethiopia, circa 2.3 Ma.19,20 KNNM-ER 1482 (mandible), KNNM-ER 60000 (mandible) and KNNM-ER 62000 (maxilla) are Kenyan specimens attributed to *H. rudolfensis* from the Turkana Basin, dated between 1.8 Ma and 2.0 Ma.21-23 OH 65 (maxilla) is a specimen of early *Homo* from Olduvai, circa 1.8 Ma.24 Early *Homo* specimens from Dmanisi in Georgia (including mandibles...
D 211, D 2600 and D 2735), circa 1.8 Ma, have been attributed to *H. erectus* or *H. georgicus*.25-26 East African specimens considered to represent *H. habilis* include the KNM-ER 1813 skull and possibly the KNM-ER 1802 mandible as part of the hominin assemblage from the Turkana Basin23,27 as well as OH 13 (mandible and maxilla) as part of the hominin assemblage from Olduvai.11,12 OH 7 is a reference specimen in this study, but for purposes of the record (Table 1), we include OH 7 MAX which is a virtual representation of the maxilla, as reconstructed by Spoor et al.8 using a technique whereby mandibular morphology can facilitate the reconstruction of a maxilla, and vice versa.

**Methods**

In a morphometric approach based on pairwise comparisons, Thackeray and Odes calculated ‘log sem’ statistics associated with regression analyses to compare OH 24 with other crania, based on measurements from original specimens published by Wood.23 The log sem results reflect variability in skull shape. In the case of the comparison between OH 24 and Sts 5 (almost complete skulls), 54 measurements are in common. In the instance of OH 24 and Sts 71 (partial skull), 44 measurements are common to both specimens. A log sem value was obtained from a comparison between our measurements of OH 7 and Sts 52b mandibles, using landmarks based mainly on points associated with mesiodistal and bucco-lingual diameters (excluding third lower molars because the specimens do not represent fully adult individuals). Forty measurements per specimen are included per specimen.

**Table 1:** Specimen numbers, taxonomic nomenclature, approximate date, skeletal element (mandible or cranium) and Z-score

<table>
<thead>
<tr>
<th>Specimen</th>
<th>Species</th>
<th>Age (Ma)</th>
<th>Skeletal element</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL 288-1</td>
<td>Australopithecus afarensis</td>
<td>Ca. 3 Ma</td>
<td>Mandible</td>
<td>1.43</td>
</tr>
<tr>
<td>AL 400-1</td>
<td>Australopithecus afarensis</td>
<td>Ca. 3 Ma</td>
<td>Mandible</td>
<td>0.71</td>
</tr>
<tr>
<td>AL 822</td>
<td>Australopithecus afarensis</td>
<td>Ca. 3 Ma</td>
<td>Mandible</td>
<td>0.47</td>
</tr>
<tr>
<td>AL 200</td>
<td>Australopithecus afarensis</td>
<td>Ca. 3 Ma</td>
<td>Mandible</td>
<td>0.71</td>
</tr>
<tr>
<td>Sts 5</td>
<td>Australopithecus africanus</td>
<td>Ca. 2.5 Ma</td>
<td>Cranium</td>
<td>1.00</td>
</tr>
<tr>
<td>Sts 71</td>
<td>Australopithecus africanus or A. prometheus</td>
<td>Ca. 2.5 Ma</td>
<td>Cranium</td>
<td>0.10</td>
</tr>
<tr>
<td>Sts 52b</td>
<td>Australopithecus africanus</td>
<td>Ca. 2.5 Ma</td>
<td>Mandible</td>
<td>1.60</td>
</tr>
<tr>
<td>AL 666-1</td>
<td>Homo sp.</td>
<td>2.3 Ma</td>
<td>Maxilla</td>
<td>3.86</td>
</tr>
<tr>
<td>OH 7</td>
<td>Homo habilis</td>
<td>1.8 Ma</td>
<td>Mandible</td>
<td>-</td>
</tr>
<tr>
<td>OH 7 MAX</td>
<td>Homo habilis</td>
<td>1.8 Ma</td>
<td>Maxilla (virtual)</td>
<td>-0.90</td>
</tr>
<tr>
<td>OH 24</td>
<td>Homo habilis</td>
<td>1.8 Ma</td>
<td>Cranium</td>
<td>-</td>
</tr>
<tr>
<td>OH 13</td>
<td>Homo habilis</td>
<td>1.8 Ma</td>
<td>Mandible</td>
<td>0.33</td>
</tr>
<tr>
<td>KNM ER 1813</td>
<td>Homo habilis</td>
<td>1.9 Ma</td>
<td>Cranium</td>
<td>1.95</td>
</tr>
<tr>
<td>KNM ER 1802</td>
<td>Homo rudolfensis or H. habilis</td>
<td>1.9 Ma</td>
<td>Mandible</td>
<td>-0.90</td>
</tr>
<tr>
<td>KNM ER 60000</td>
<td>Homo rudolfensis</td>
<td>1.9 Ma</td>
<td>Mandible</td>
<td>2.90</td>
</tr>
<tr>
<td>KNM ER 1482</td>
<td>Homo rudolfensis</td>
<td>2.0 Ma</td>
<td>Mandible</td>
<td>4.33</td>
</tr>
<tr>
<td>KNM ER 62000</td>
<td>Homo rudolfensis</td>
<td>2.0 Ma</td>
<td>Partial face</td>
<td>5.29</td>
</tr>
<tr>
<td>OH 65</td>
<td>Homo rudolfensis</td>
<td>1.8 Ma</td>
<td>Maxilla</td>
<td>4.81</td>
</tr>
<tr>
<td>D 211</td>
<td>Homo erectus</td>
<td>1.8 Ma</td>
<td>Mandible</td>
<td>1.95</td>
</tr>
<tr>
<td>D 2600</td>
<td>Homo erectus</td>
<td>1.8 Ma</td>
<td>Mandible</td>
<td>2.90</td>
</tr>
<tr>
<td>D 2735</td>
<td>Homo erectus</td>
<td>1.8 Ma</td>
<td>Mandible</td>
<td>2.43</td>
</tr>
</tbody>
</table>
were taken from virtual reconstructions of OH 7 and Sts 52b by Spoor et al. and Benazzi et al. respectively. Sts 52b was selected for comparison in this study because of an apparent degree of morphological similarity with OH 7, as discussed by Tattersall. Using Procrustes Distances (PD), Spoor et al. compared Plio-Pleistocene hominin specimens with a focus on OH 7, based on more than 50 landmarks. For purposes of our study, we used PD data made available through Spoor and Gunz (personal communication to JFT, 2020). Here we integrate the two kinds of shape-related statistics (PD and log sem) by expressing them as standardised Z-scores in relation to data obtained from humans and extant great apes.

The AL 400-1 mandible of A. afarensis has been compared with OH 7 on the basis of the PD method. AL 400-1 has also been compared with OH 7 using the log sem statistic. The difference between Z-scores is expected to be relatively small if the approaches yield consistent results.

Log sem statistics for crania

The log sem statistic has been previously used in analyses of linear measurements obtained from crania of modern specimens in natural history museums and Plio-Pleistocene hominins. In this method, measurements are subjected to pairwise comparisons, using least squares linear regression to generate an equation of the form \( y = mx + c \), where \( m \) is the slope and \( c \) is the intercept. In an initial study of pairs of specimens of the same (extant) species in 1997, Thackeray et al. reported central tendency of the log-transformed standard error of the m-coefficient, known as ‘log sem’ which is a measure of the degree of scatter around the regression line, reflecting variability in shape. Central tendency of log sem was also discovered using larger samples, associated with a mean log sem value of -1.61 reported in 2007 by Thackeray. At least for hominoids, the mean log sem value of -1.61 ± 0.1 was recognised in 2016 as a typical degree of intraspecific variation. Two log sem values are obtained. The absolute difference between these values is termed ‘delta log sem’. The mean delta log sem is small for intraspecific comparisons, whereas the delta log sem values calculated for interspecific comparisons are generally larger.

Research Article

In response to views expressed by Gordon and Wood, Thackeray and Dykes stated that the number of measurable dimensions (k) in interspecific comparisons is expected to be relatively small if the approaches yield consistent results. A criticism that has been levelled against the log sem approach relates to which variables are being measured. Remarkable as it may seem, the degree of intraspecific variability reflected by a mean log sem value of -1.61 ± 0.1 has been obtained, not only from cranial variables, but also from measurements from teeth.

Procrustes Distances for mandibles

Bookstein and Duta have described the method whereby PD are calculated, reflecting differences in shape between objects (in this case, mandibles). PD statistics serve to quantify the difference between landmarks by using the square root of the sum of squared differences in positions of those landmarks.

Spoor et al. calculated PD values from pairwise comparisons of OH 7 and other hominins, using landmarks indicated in their Fig. 2f and 2g. PD were also calculated for purposes of comparisons with H. sapiens, P. troglodytes and G. gorilla. The mean PD for the extant hominoids, based on data obtained by Spoor and Gunz (personal communication to JFT, 2020), provides a ‘within group’ (conspecific) frame of reference.

Z-scores

The mean and standard deviation for PD for pairwise comparisons of extant hominoids (0.089 ± 0.021) are analogous to the mean and standard deviation for log sem values for extant hominoids (-1.61 ± 0.1). These two means can both be related to probabilistic Z-scores, where the Z-value of 0 corresponds to the mean value of 0.089 (PD) and also to the mean log sem value of -1.61. One standard deviation above or below the mean is (approximately) 1. These Z-scores are expected to be relatively small if the approaches yield consistent results. Likewise, two standard deviations above or below a mean Z-value of 0 correspond to Z-score values of greater than 20, there is a tendency for the mean log sem to stabilise around a value of circa -1.6.

In their analyses of cranial measurements of Pan troglodytes, P. paniscus, Gorilla gorilla and H. sapiens (using more than 20 measurable dimensions as published by Gordon and Wood), Thackeray and Dykes obtained the following results from pairwise comparisons:

1. Mean log sem = -1.61 ± 0.129 (n = 8072 pairwise regressions) reflects what is considered to be a typical degree of intraspecific variation within hominoids.

2. Mean log sem = -1.62 ± 0.126 (n = 26780 regressions), reflecting the degree of variability from interspecific comparisons.

These results, based on a very large number of regressions, clearly show that there is indeed a significant difference between the log sem values calculated for intraspecific and interspecific comparisons, which is related to similarity (or dissimilarity) in shape.

Figure 2: Visualisation of Z-scores for hominins based on data in Table 1. OH 7 and OH 24 are the reference specimens for pairwise comparisons with other hominins.

Z-scores and Plio-Pleistocene hominins

Page 3 of 5
citra 2 and -2, respectively. In this way, log sem statistics (including those calculated for a comparison between OH 24 and skulls of other hominins) and PD values (including those calculated for comparisons between OH 7 and other hominin mandibles) can be expressed on a common (2) scale.

Results

Results from this study are presented in Table 1. Figure 2 is a visualisation of the results without a time scale.

The AL 400-1 mandible of *A. afranensis* has a Z-value of 0.71 when it is compared with OH 7 using the PD statistical method. A Z-value of 1.0 is obtained when AL 400-1 is compared to OH 7 using the log sem statistical method. The difference of only 0.29 is relatively small, reflecting consistency.

As indicated in Table 1 and Figure 2, the Z-score for the Sts 52b jaw (compared to the OH 7 mandible) is similar to that which has been obtained for the Sts 5 skull (compared to the OH 24 cranium), and similar also to the Z-score for the Sts 71 skull (also compared to OH 24).

In Figure 2 there is a clear separation between specimens attributed to *H. habilis* (including KNM-ER 1813 and OH 13) and others attributed to *H. rudolfensis* (e.g. KNM-ER 62k and KNM-ER 1482). In terms of Z-scores, specimens from Dmanisi attributed to *H. erectus* are intermediate between *H. habilis* and *H. rudolfensis*.

Discussion and conclusions

We integrated two statistics (PD and log sem) by expressing them as standardised Z-scores. As the important type specimen of *H. habilis*, the OH 7 mandible has been used as a frame of reference for quantifying PD between it and other hominin mandibles. The contemporary OH 24 skull has been used for purposes of calculating log sem statistics in the context of pairwise comparisons between it and other hominin skulls (notably Sts 5 and Sts 71) from South Africa. With OH 7 and OH 24 representing the same species, measurements of mandibles and skulls have been used to obtain PD and log sem statistics, respectively, expressed on a common Z-score scale relative to values for conspecific extant hominoids.

There is a wide spectrum of variation in Z-scores for Tanzanian and Kenyan specimens attributed to early *Homo* dated within the period between circa 1.8 Ma and 2 Ma. By contrast, in terms of variation through time predating 1.8 Ma, Z-scores for *A. afranensis*, *A. africanus*, and *H. habilis* display a relatively small range of variability. Put another way, the Sterkfontein specimens are a similar morphological distance to the Olduvai hominins using Z-score measurements. The Z-scores of the specimens from Hadar fall within the range of those determined for their reconstruction of OH 7 compared to other hominins. We are grateful to Leslie Aiello, José Braga, Marine Cazeneve, the late Sue Dykes, Jean-Michel Loubes, Cattlin Schrein, Clément Zanolli and anonymous referees for encouragement and advice in the course of preparation of this article.

Competing interests

We have no competing interests to declare.

Authors’ contributions


References


