Identifying bilateral pairs of deer (*Odocoileus* sp.) bones: how symmetrical is symmetrical enough?

R. Lee Lyman

Department of Anthropology, University of Missouri—Columbia, 107 Swallow Hall, Columbia, MO 65211, USA

Received 9 November 2005; received in revised form 22 December 2005; accepted 3 January 2006

Abstract

Identifying bilaterally paired bones in zooarchaeological collections rests on the assumption that left and right skeletal elements from the same organism will be symmetrical. Bilaterally paired bones in an individual are, however, typically asymmetrical to some degree, demanding that the question “How symmetrical is symmetrical enough to identify a bilateral pair?” be answered with control data. The degree of symmetry chosen, or tolerance, will influence both how many true pairs in an archaeological collection are not identified (type I error) and how many false pairs are identified (type II error). Bivariate measures of 60 pairs of astragali and 48 pairs of distal humeri of deer (*Odocoileus virginianus* and *O. hemionus*) indicate that both sorts of error are frequent even when a conservative level of tolerance (average asymmetry) is used. Simulation of an archaeological collection indicates that as sample size increases, frequencies of both kinds of error increase. Application of the matching criteria and tolerance level to an archaeological collection underscores that the analytical requirements of identifying paired bones are steep. © 2006 Elsevier Ltd. All rights reserved.

Keywords: Anatomical refitting; Bilateral symmetry; Deer

1. Introduction

Various analytical techniques in zooarchaeology require that bilaterally paired bones such as left and right femora, left and right humeri, and the like be identified among commingled left and right skeletal elements representing multiple individuals. These include estimation of the number of individual animals represented using the Lincoln index (e.g., [1,5,11,35,38]), and measuring the dispersal of the bones of individual carcasses to detect natural site formation processes or hominid behaviors involving butchering and food sharing (e.g., [3,8–10,32,37]).

Analytical identification of bilaterally paired bones rests on the assumption that the members of a pair—the left and right mates—are bilaterally symmetrical. Members of bilateral pairs are, however, typically asymmetrical to some degree. It is therefore critical to determine how asymmetrical paired bones are in order to conclude that two bones that might be paired are indeed paired, or are from the same individual animal. Some have argued that determination of how symmetrical is symmetrical enough is a slippery issue because symmetry varies within a species from nearly perfect to somewhat asymmetric [12]. Indeed, temporally fluctuating asymmetry has in the past few decades become a very important research topic in biology ([18,28] and references therein) in which metrical size bilateral symmetry is usually measured [27].

Early on, some paleozoologists suggested that it is fairly straightforward to establish how symmetrical is symmetrical enough to identify a bilateral pair (e.g., [17]). The general procedure is to establish the range of variation in (a)symmetry for particular skeletal elements within a species (e.g., [10,13,24,32,33]). But the specific procedure followed when deciding how similar is similar enough to conclude that a bilateral pair has been identified among archaeological specimens is typically obscure (e.g., [4,10,13,24,32,33,37]).

Nichol and Creak [25, p. 15] state that the more possible pairs in an archaeological collection, the more likely that pairs will be found and the smaller the “error margin needed to
allow all the bones from one side to be accommodated by those from the other.” This assumes that every left element will have a matching right element within a collection. This assumption may be reasonable when one searches for bilateral pairs among skeletal elements that were accumulated and deposited together, such as individual skeletal elements that make up a large portion of the skeleton and that can be assumed to have been regularly accumulated as a unit. An example is the numerous individual elements that comprise a fish head, precisely the elements that Nichol and Creak [25] examined. It seems progressively less reasonable to make this assumption as the size of the animal increases and as the natural anatomical propinquity of the paired elements decreases. If the elements of a bilateral pair are large, and are anatomically not near each other within a body, such as with the humeri of an ungulate carcass too large to be transported as a complete unit, then they are easily disassociated anatomically and are potentially accumulated and deposited independently of one another. If the elements are relatively small and are anatomically interdependent (articulated or within the same anatomical region), such as with left and right mandibles of a rabbit, then they are likely to be accumulated and deposited interdependently.

Nichol and Creak [25, p. 8] referred to the maximum allowable difference (or degree of asymmetry) between bilateral pair mates as the “tolerance.” Tolerance should be empirically determined from a set of known reference specimens, but in doing so, several things must be acknowledged. The tolerance level will depend on the attributes chosen. Genetic and environmental conditions influence the degree of asymmetry in anatomical attributes differentially [18]. Further, different populations will display different degrees of asymmetry in the same attribute [18]. Thus the degree of symmetry necessary—the tolerance—to identify pairs determined on the basis of a modern reference collection will be to some unknown degree specific to that collection. The average degree of asymmetry in an attribute displayed by a modern population may or may not be equal to the average asymmetry manifest in a prehistoric population.

Another thing requiring acknowledgment is the analyst’s willingness to make identification errors. Regardless of the chosen tolerance level, a left and a right element might be misidentified as not a pair when in fact they are a true pair (type I error), and a match might be identified when in fact the left and right elements do not represent a true pair (type II error). A tolerance level that minimizes both type I and type II errors in matching must be empirically established by using museum skeletons, regardless of the anatomical criteria—metric (size), morphological (shape), meristic (frequency)—used to identify bilaterally paired skeletal elements. This simple yet central point of all bilateral refitting methods is illustrated in the remainder of this paper.

2. Materials and methods

The astragalus and the distal humerus are two compact, dense skeletal parts that survive in measurable condition despite many sorts of taphonomic processes [6,19,20]. By choosing two skeletal parts, I hoped to avoid the problem of using one among the specimens of which it might be difficult to identify bilateral pairs. Results for astragali and distal humeri are very similar and strengthen the conclusions drawn.

Deer (Odocoileus spp.) are widely distributed prehistorically across much of the Americas, their remains are frequent in archaeological contexts, and numerous modern skeletons are housed in museums. I measured museum specimens of two species of deer—white-tailed deer (O. virginianus) and mule deer (O. hemionus)—collected from the western United States. A total of 60 pairs of astragali (left and right elements) from adult (all epiphyses fused or fusing, skeletal elements of adult size) museum specimens was measured; 17 pairs are of white-tailed deer, and 43 pairs are of mule deer. A total of 48 pairs of distal humeri from some of the same adult museum skeletons was measured; 17 pairs of white-tailed deer, 30 pairs of mule deer, and 1 hybrid of the two species. Based on recent research both skeletal parts can now often be distinguished as to species based on morphological features [14,15], and these features were used to taxonomically distinguish specimens in the archaeological collection [21].

I focus here on linear measurements rather than, say, rugosity of muscle attachment scars and number of foramina (e.g., [29]). Two measurements were taken on each astragalus and each distal humerus—both museum specimens and archaeological specimens—to the nearest 0.02 mm with sliding calipers. The two measurements of astragali were the (greatest) distal breadth (Db [7]) and the (greatest) lateral length (GLl [7]). The two measurements of distal humeri were the anterior breadth of the distal trochlea (DBt) and the minimum (antero-posterior) diameter of the (latero-medial) center of the distal trochlea (MNd). Some archaeological specimens of astragali and of distal humeri that could be identified to species were not measurable and some specimens that were measured could not be identified to species; various specimens were broken or weathered and lacked taxonomically diagnostic features; others displayed taxonomically ambiguous features. Thus the frequencies of specimens identified to species reported here are different than those reported elsewhere [21].

Use of two measurements rather than one should make any effort to identify bilateral mates more accurate because specimens must be relatively symmetrical in two dimensions rather than just one. Given how the two measures are analyzed (see below), specimens could be assessed for symmetry of shape. Minimizing measurement error is critical to studies of asymmetry [27]. I measured a sample of astragali (5 lefts, 5 rights, no bilateral pairs) once each day over a period of 3 days. The average difference in measurements approximates measurement error. The tolerance level should exceed the measurement error to insure true pairs are not overlooked because of measurement error. Measurement error of astragalus distal breadth was 0.062 ± 0.06 mm, and of greatest lateral length was 0.073 ± 0.05 mm.

Measurements of all 60 astragalus pairs (Table 1) and measurements of all 48 pairs of distal humeri were compiled (Table 2). To establish how symmetrical was symmetrical...
enough in a collection of known pairs, conceive of the absolute value of the difference between the two measurements (on left and right mates) of one dimension (whether of astragali or distal humeri) as defining the length of one side of the right angle of a right triangle (side a in Fig. 1, upper), and the absolute difference between the two measurements (on left and right mates) of the other dimension as defining the length of the other side of the right angle of the right triangle (side b in Fig. 1, upper). The Pythagorean theorem \(a^2 + b^2 = c^2\) is used to find how asymmetrical the left and right specimens are in terms of the two dimensions (Fig. 1,
lower). If the left and right specimens are perfectly symmetrical, then the hypotenuse (side $c$) of the right triangle will be zero because $a = 0$ and $b = 0$. What is hereafter referred to as the $c$-value (the length of the hypotenuse) is, then, a measure of (a)symmetry. The $c$-value by itself suggests symmetry in terms of size. A graph of the right triangles formed by left and right mates suggests symmetry in terms of shape because both dimensions are considered simultaneously. I nevertheless focus on size.

There was no statistically significant difference between the astragalus mean $c$-values of the two species (white-tailed deer, $0.315 \pm 0.164$; mule deer, $0.359 \pm 0.217$; Student’s $t = 0.757; p = 0.45$), so the mean of all 60 pairs was determined ($0.347 \pm 0.203$). This grand mean $c$-value indicates how symmetrical astragalii are in deer in terms of the two dimensions measured. Using both species as the basis for a tolerance level, most conservatively, any left astragalus of a species that produces a $c$-value $\leq$ the grand mean $c$-value (0.347) with a right astragalus of the same species could be a match. This value is the liberal matching statistic. Applying both the conservative and the liberal matching statistics to known pairs from which they were derived will reveal how varying tolerance levels influence the magnitude of type I and type II errors.

There was no statistically significant difference between mean $c$-values for distal humeri of the two deer species (white-tailed deer, $0.753 \pm 0.81$; mule deer, $0.467 \pm 0.282$; Student’s $t = 1.771; p > 0.08$), so the grand mean was determined; $c$-value $= 0.561 \pm 0.541$. The conservative tolerance level or matching statistic is ($c$-value $= 0.561$; the liberal matching statistic is 1.102.

The archaeological collection originated in the site of Cathlapotle (45CL1), in southwestern Washington State [2,21,23]. The site was visited by Meriwether Lewis and William Clark in March of 1806 as they led the Corps of Discovery eastward. At that time the site comprised several large cedar-plank houses and associated midden deposits [2]. Radiocarbon dates indicate the main occupation began about AD 1450. Ceramic trade goods indicate abandonment circa AD 1834. I identified all mammalian remains. Site deposits were designated as exterior or interior deposits—outside or inside of a house, respectively. Exterior “middens” deposits have very high organic contact, lenses of fresh-water mussel shells, and other indications that they formed as primary or secondary dumps [2]. Exterior yard deposits are generally broad sheet-like deposits that contain intact hearths, activity areas, pits, evidence of small structures, and so forth. They lack the very high organic content of middens. Interior deposits were variously assigned to walls, benches (deposits below the 2 m wide sleeping benches or platforms that ran along the interior side of the house walls), storage pits, and hearth areas. The houses had extensive subfloor storage features that are about 2 m wide by 2 m deep located below the sleeping platforms that line the interior walls of the houses.

Most mammalian remains from Cathlapotle were recovered from storage pits and exterior midden deposits. Site deposits are distributed over an area about $200 \times 50$ m. Less than 2% of this area has been excavated [2]. No change in the size of the deer bones over time is evident [22] so I treat the deer remains as a single assemblage. Metric and general provenience data for deer astragalii from Cathlapotle are described in Table 3.

3. Results using museum specimens

3.1. Initial results by species

The bivariate scatterplot of the 17 white-tailed deer astragalus pairs (Fig. 2) from museum collections indicates that simple visual inspection of this graph, were it based on an archaeological collection, might provide sufficient resolution to allow some astragali to be correctly paired. But there is a cluster of three left and three right astragalii, and another cluster of four left and four right astragalii within each of
which it would be difficult to identify true pairs. The conservative pair statistic does not help identify pairs in either cluster. Both clusters are near the center of the plot. The GLI and Db values of all deer are normally distributed (do not differ significantly from a Gaussian distribution; $p > 0.1$). It is in the center of a distribution of points (Fig. 2) where type II errors (false pairs identified) would be most likely to occur and these would also likely influence the magnitude of type I error (e.g., if a false pair is identified and the true mate of one of the members of that pair occurs in the assemblage).

Both probability theory and the empirical data in Fig. 2 indicate that true pairs will be most obvious at the extremes of the point scatter (upper right, lower left) where few specimens occur in normally distributed variables. But at these extremes the conservative pair statistic would not identify three of the seven pairs plotted (all in the upper right portion of the graph) if only that statistic was used to identify pairs. In this case, the graph rather than the statistic helps avoid type I error (failure to identify a true pair). This is likely because in this case, the conservative pair statistic indicates that the left and right members of a pair must display a $c$-value $\leq 0.366$; the liberal pair statistic indicates that the left and right members of a pair must display a $c$-value $\leq 0.563$ ($=0.366 + 0.197$). But that is all those statistics indicate. They do not indicate where in a population (as represented by all the points in Fig. 2) any given possible pair falls.

The bivariate scatterplot of the 43 mule deer astragalus pairs (Fig. 3) from museums indicates that with a sample of
Study of bilateral pairs of astragali of each species suggests that as the absolute size of the archaeological sample of possible pairs increases, the probability of committing some type II error (and related type I error) will increase as well. As more possibly matching archaeological specimens are included in an analysis, it will be much more difficult to identify true pairs. The number of matches will increase only if the distinction of type I and type II errors is ignored. Earlier analysts seem to have recognized this problem regarding large samples, though their wording was inexplicit (e.g., [36, p. 311]). Later commentators were more explicit [11]. Nichol and Wild [26, p. 36–37], for instance, noted that

"it is much harder to identify the extra unmatchable elements in a collection of bones from a larger number of animals of a species than it is in a smaller collection. For example, if there are 100 lefts and 100 rights of a bone, it will be rather difficult to produce an estimate of MNI much greater than 100, whether 100 or 200 animals are represented, while it may be very easy to see that a single left and a single right come from different individuals."

Consider Fig. 4, which shows the 48 pairs of museum distal humeri from some of the same deer that provided the astragali data. Notice that the points representing three lefts overlap the points representing three rights in Fig. 4. But none of the true pairs are actually symmetrical (Table 2). The three pairs of overlapping points all represent false pairs. And, the tight clustering of points in the middle of the graph suggests that other false pairs would likely be identified even were the conservative matching statistic used.

### 3.2. Simulating an archaeological collection

Simulating an archaeological collection using a set of known pairs allows evaluation of how many instances of type I error and how many instances of type II error occur when different tolerance levels are used. I drew a random sample of 60 astragali (without replacement) from among the 120 comprising the collection of museum specimens. (Astragali were numbered from 1 through 120 in the order listed in Table 1. Members of the first pair of astragali are numbered,
respectively, 1 and 2, members of the last pair are 119 and 120; right astragali all have odd numbers, left astragali all have even numbers.) I initially ignored species designations because until a few years ago the two deer species could not be distinguished among zooarchaeological remains and, as noted earlier, some astragali cannot be identified to species even today. The sample of 60 astragali included 28 left elements and 32 right elements, meaning there were 28 possible pairs; 16 true pairs were included. The bivariate scatterplot of the sample (Fig. 5) reveals many of the problems with metrically identifying bilateral pairs, even with explicit tolerance limits.

Of the 16 true pairs, only four (astragali 7–8, 13–14, 59–60, 83–84) would be identified if the conservative matching statistic were used ($c$-value = 0.347). But use of that conservative statistic would also result in the identification of seven false pairs (5–98, 6–97, 20–79, 30–63, 32–93, 36–73, 104–119). Using the liberal matching statistic would add three true pairs (55–56, 71–72, 99–100). The conservative matching statistic creates a plethora of problems many of which would be exacerbated by the liberal matching statistic used. The conservative matching statistic results in ignoring six true pairs (5–6, 19–20, 35–36, 73–74, 79–80, 97–98). The conservative matching statistic does not help sort out the cluster of points in the lower left portion of the graph (Fig. 5). That cluster consists of several true pairs (1–2, 13–14, 59–60, 91–92), only one (59–60) or two (59–60, 91–92) of which would be identified depending on whether or not one seeks to identify the maximum possible number of pairs. One possible solution concerning this point cluster would identify three pairs (17–92, 91–102, 59–60), only one of which was a true pair; another solution would identify four pairs (17–78, 91–92, 1–102, 59–60), only two of which were true pairs.

If species are recognized among the astragali, fewer errors occur but a significant number remain. Five of the seven false pairs would not be identified, but two would still be identified. All 6 true pairs not recognized by the conservative matching statistic would still not be identified. The cluster of points in the lower left of Fig. 5 would sort out better, but at least one false pair would remain (91–102) and a true pair (1–2) would be missed if the conservative matching statistic were used. As noted earlier, and as can be imagined by inspection of Fig. 4, similarly messy results are produced by the distal humerus data.

4. Results with an archaeological collection

Does the difficulty of identifying bilateral pairs attend study of a real archaeological collection? The mule deer astragali and the deer of unknown species (Odocoileus sp.) astragali from Cathlapotle are plotted in Fig. 6. The latter are those that did not clearly display morphological criteria diagnostic of either species, thus they may represent either species. Using the conservative matching statistic, there seems to be one pair of left and right astragali in this set—specimens 1 and 3 ($c$-value = 0.2). Both members of the pair clearly represent mule deer based on morphological features. There are no pairs among the specimens identified only as deer using either the conservative or the liberal matching statistic for mule deer. None of the specimens identified only to genus pair up with the specimens identified as mule deer, thus the former will be compared to those astragali representing white-tailed deer.

The white-tailed deer and the deer astragali from Cathlapotle are plotted in Fig. 7. As noted earlier, the latter are included because they may represent white-tailed deer although they did not display the diagnostic morphology of either species represented in the collection. Using the conservative matching statistic for white-tailed deer ($c$-value = 0.315), there are two pairs of left and right astragali; specimen 9 matches with specimen 22 ($c$-value = 0.072), and specimen 10 matches with specimen 46 ($c$-value = 0.321). The liberal matching statistic ($c$-value = 0.479) suggests there may be two more pairs, specimens 21 and 33 ($c$-value = 0.362), and specimens 8 and 16 ($c$-value = 0.439).

Fig. 5. Simulated archaeological assemblage of astragali. Each number signifies a particular astragalus listed in Table 1; odd numbers are right astragali, even numbers are left astragali (see text for explanation). Some numbers have been moved slightly to enhance legibility.

Fig. 6. Cathlapotle mule deer (O. hemionus) and deer (Odocoileus sp.) astragali. The single bilateral pair identified using the conservative matching statistic is circled with a dashed line, and specimen numbers (Table 3) are included.
1. Among the deer remains at Cathlapotle we have not gained any resolution by identifying bilateral pairs. Nomic abundances are at best rank-order or ordinal scale [12], of skeletal elements of both bighorn sheep (Ovis canadensis) and mule deer and 154 for white-tailed deer [21]. Given that taxonomic abundances are rank-order or ordinal, about the same distance apart [33]. The conjoined fragments of skeletal elements of both bighorn sheep (Ovis canadensis) and mule deer, there are 24 lefts, 16 rights, and 3 pairs. Plugging these values into the Lincoln index of (LR)/P, where L is the number of lefts, R is the number of rights, and P is the number of pairs [11,35], estimates of 45 individual mule deer and 128 individual white-tailed deer are derived. The estimated number of individual mule deer is less than the estimated number of individual white-tailed deer. The number of identified specimens (NISP) of all bones identified for these two species is 23 for mule deer and 154 for white-tailed deer [21]. Given that taxonomic abundances are at best rank-order or ordinal scale [12], we have not gained any resolution by identifying bilateral pairs among the deer remains at Cathlapotle.

The mean distance between bilaterally matched bison (Bison antiquus) long bones at the Horner bison kill site is about 1–2 m, and conjoined fragments of skeletal elements are about the same distance apart [33]. The conjoined fragments of skeletal elements of both bighorn sheep (Ovis canadensis) and bison (Bison bison) at the Bugas-Holding site average ≤1 m apart [30]. At the Jones-Miller bison (Bison sp.) kill site, conjoined fragments average about 2 m apart [34].

Based on Waguespack’s [37, p. 411] map and tabular summary of refits at a habitation site, I estimated the distance between 17 bilateral refits. For those that seemed to occupy the same horizontal position (n = 7), I assumed specimens making up a pair were 0.25 m apart. The 17 bilateral refits are, on average, about 4.9 m apart. This is more than twice as far apart as bilaterally refitting bones in kill sites. Enloe [9, p. 15] found portions of a single carcass to be separated by as much as 63 m in the Upper Paleolithic habitation site of Pincevent. Thomas’s [31, p. 367] 35-year-old assumption that “the dietary practices of man tend to destroy and disperse the bones of his prey-species” is confirmed and expanded. Bones of a carcass are likely to be more dispersed (farther apart) within a habitation site than the bones of a carcass of that taxon at a kil or procurement site.

But, consider each of the five pairs identified at Cathlapotle. According to the excavator, K.A. Ames, both members of the 1–3 pair (Table 3) come from a sheet midden; the specimens are about one meter apart horizontally and could well represent a true pair. Specimen 10 comes from the single unit excavated inside House 2 and its mate, specimen 46, comes from a sheet midden outside of House 2 such that the two specimens were 20 m apart; Ames suggests the probability that these specimens represent a true pair is remote. Specimen 9 comes from House 2 and specimen 22 from House 4; they were about 21 m apart and if a true pair, then represent interhouse sharing. Both specimens of the 8–16 pair came from midden deposits and were 55 m apart; there is no clear evidence of post-depositional disturbance so they may represent a true pair. Finally, specimen 21 came from the floor of House 4 and its mate, specimen 33, came from a storage pit inside of House 1; the two specimens were 60 m apart and likely represent a false pair. Given the absence of clear ethnoarchaeological indications of the amount of dispersal to expect at habitation sites, the lack of evidence corroborating suspected pairs, and the difficulty of identifying pairs exemplified here, I am hesitant to say much else about these five pairs.

6. Conclusion

Chaplin [4, p. 70] noted that “the right and left bones of an animal are not always the same size.” Winder [38, p. 116] indicated that “the reliable recognition of [bilateral] pairs in real assemblages may be difficult.” Fieller and Turner [11, p. 57] emphasized that identifying pairs could present an “immense problem” if the sample of possible pairs was large. I have shown that the sample size need not be very large before the immensity of the matching problem arises. A mere 15 possible pairs presents what may be insurmountable difficulties, in part because the rules for determining how similar is similar enough are analytical choices. Consideration of two dimensions simultaneously and use of a tolerance level in the form of a conservative matching statistic does not make things easier. Nor do these aspects of analytical protocol reduce type I or type II error.
Many researchers cited here use metric symmetry to provide an initial indication of possible pairs. Enloe and David [10, p. 297] found that size variables identified multiple possible mates for individual bones, and that confirmation of a pair’s validity depended on “morphological symmetry, architecture of the bones, and the development of muscle and ligament attachment points.” With large samples made up of several dozen specimens of each side of multiple skeletal elements, the requisite metric data become numerous, the time necessary to visually compare morphologies becomes long, laboratory layout space requirements increase, and the like. Biologists are developing multivariate techniques for comparing shapes of paired bones (e.g., [16]) that demand statistical sophistication. Any effort to identify bilaterally paired bones, regardless of the criteria and methods used, must be attended by explicit recognition of the necessity of simultaneously reducing both type I and type II error. In light of the analyses described above, future identifications of bilateral pairs should be attended by deep data and analytical requirements.

Acknowledgments

Excavations at Cathlapotle were supported by the U.S. Fish and Wildlife Service and the Friends of the Wapato Valley. Anan Raymond, Regional Archaeologist for the U.S.F.W.S. and his staff, the staff of the Ridgefield Wildlife Refuge, and the people of Ridgefield, Washington, supported excavations in diverse ways. The Chinook Tribe also provided support and keen interest, and actively participated in the project. Analysis of the Cathlapotle mammalian fauna was partially funded by the College of Arts and Science of the University of Missouri—Columbia. Support for the collection of metric data was provided by the Chancellor’s Office, University of Missouri—Columbia, and Barbara M. Lyman. Kristen Fuld suggested the term “c-value.” Helpful comments on an early draft were provided by D.K. Grayson and two anonymous reviewers.

References


