Humeral Head Circle-Fit Method Greatly Increases Reliability and Accuracy When Measuring Anterior–Posterior Radiographs of the Proximal Humerus

Chad S. Mears,[1] Tanner D. Langston,[1] Colton M. Phippen,[1] Wayne Z. Burkhead,[2] John G. Skedros[1]

[1]Department of Orthopaedics and Utah Orthopaedic Specialists, University of Utah, 5323 South Woodrow Street, Suite 200, Salt Lake City 84107 Utah, [2]The W.B. Carrell Memorial Clinic, Dallas, Texas

Received 11 July 2016; accepted 28 December 2016
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jor.23520

ABSTRACT: Measurements made on routine A–P radiographs can predict strength/quality of the proximal humerus, as shown in terms of two easy-to-measure parameters: Cortical index (CI) and mean-combined cortical thickness (MCCT). Because of high variability inherent when using established methods to measure these parameters, we describe a new orientation system. Using digitized radiographs of 33 adult proximal humeri, five observers measured anatomical reference locations in accordance with: (i) Tingart et al. (2003) method, (ii) Mather et al. (2013) method, and (iii) our new humeral head Circle-Fit method (CFM). The Tingart and Mather methods measure CI and MCCT with respect to upper and lower edges of 20 mm tall rectangles fit to a proximal diaphyseal location where endosteal (Tingart) or periosteal (Mather) cortical margins become parallel. But high intra- and inter-observer variability occurs when placing the rectangles because of uncertainty in identifying cortical parallelism. With the CFM an adjustable circle is fit to the humeral head articular surface, which reliably and easily establishes a proximal metaphyseal landmark (M1) at the surgical neck. Distal locations are then designated at successive 10 mm increments below M1, including a second metaphyseal landmark (M2) followed by diaphyseal (D) locations (D1, D2 $\cdots$D6). D1 corresponds most closely to the proximal edges of the rectangles used in the other methods. Results showed minimal inter-observer variations (mean error, 1.5 ± 1.1 mm) when the CFM is used to establish diaphyseal locations for making CI and MCCT measurements when compared to each of the other methods (mean error range, 10.7 ± 5.9 to 13.3 ± 6.7 mm) ($p < 0.001$). © 2017 Orthopaedic Research Society. Published by Wiley Periodicals, Inc. J Orthop Res

Keywords: humerus; radiograph; cortical index; circle-fit method; reliability; cortical thickness

Morphological measurements made using routine anterior–posterior (A–P) radiographs of the proximal humerus are becoming more common for clinical and research purposes as a way to predict bone strength/quality.[1–4] For example, cortical index (CI) is the most common measurement employed in clinical settings.[1,5–9] CI is defined as the difference between the outer (OD) and inner diameters (ID) divided by the OD [(OD-ID)/OD] (lower CI values = weaker bone).[6,10] Mean combined cortical thickness (MCCT = (OD-ID)), which is obtained from two levels of the proximal humerus diaphysis (hence "combines" four cortical thickness measurements), is a related measurement that has been shown to even more strongly correlate with bone strength.[1,5,7,11] CI and MCCT are clinically relevant: (i) surgeons can evaluate radiographs of the fractured humerus and the non-fractured side and make decisions regarding choices for techniques used for surgical reconstruction,[2,3,12] (ii) age-related changes in these simple measurements correlate with reduced bone quality and fracture strength of the proximal humerus[1,5,8] and other bones (see overview in,[13]) and (iii) age-related changes can result in complications of shoulder arthroplasty and fracture fixation.[3,6]

One of the more popular methods for making CI and MCCT measurements on A–P radiographs of the proximal humerus diaphysis is that of Tingart et al. (2003)[5] (Fig. 1). This method is based on fitting a 20 mm tall rectangle to the proximal diaphysis where the endosteal cortical margins become parallel (i.e., the proximal-most [upper] edge of the rectangle is placed where parallelism starts, as seen when viewing the endosteal margins from proximal to distal). While this method can be easily done using A–P radiographs obtained routinely in the clinical setting,[2,3] our pilot studies revealed that this method incurs high inter- and intra-observer errors because of the uncertainty in defining parallelism of the endosteal cortical margins.[14]

Mather et al.[11] used a modified version of the Tingart method. This modification is based on fitting a 20 mm tall rectangle to the location where the outer cortical (periosteal) margins become parallel. Similarly, our pilot studies also showed high inter- and intra-observer errors when using the Mather method.

Our pilot studies also showed that the errors in determining the reference locations using these two methods are highly clinically relevant. This is because the error range resulted in significant differences in the correlations of bone strength of mechanically tested humeri where these methods were used to measure CI and MCCT prior to fracturing the bones.[14] These findings bespeak the need for improved anatomical reference/orientation methods for making CI, MCCT, and related measurements on A–P radiographs of the proximal humerus.

The purpose of this study is to introduce a simple method that can be performed in clinical settings that is more reproducible for making CI and MCCT measurements than the Tingart, Mather and related methods.

Conflicts of interest: None.

Correspondence to: John G. Skedros, (T: 801-747-1020; F: 801-747-1023; E-mail: jskedrosmd@uosmd.com)

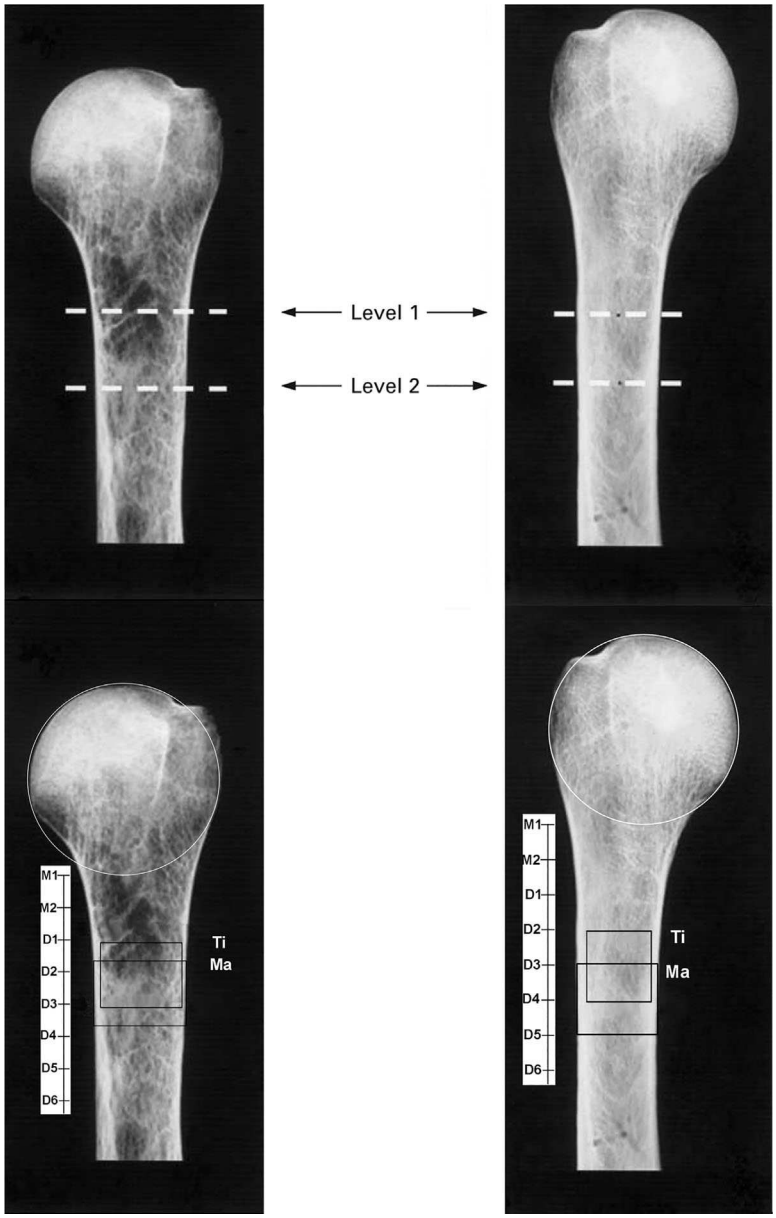
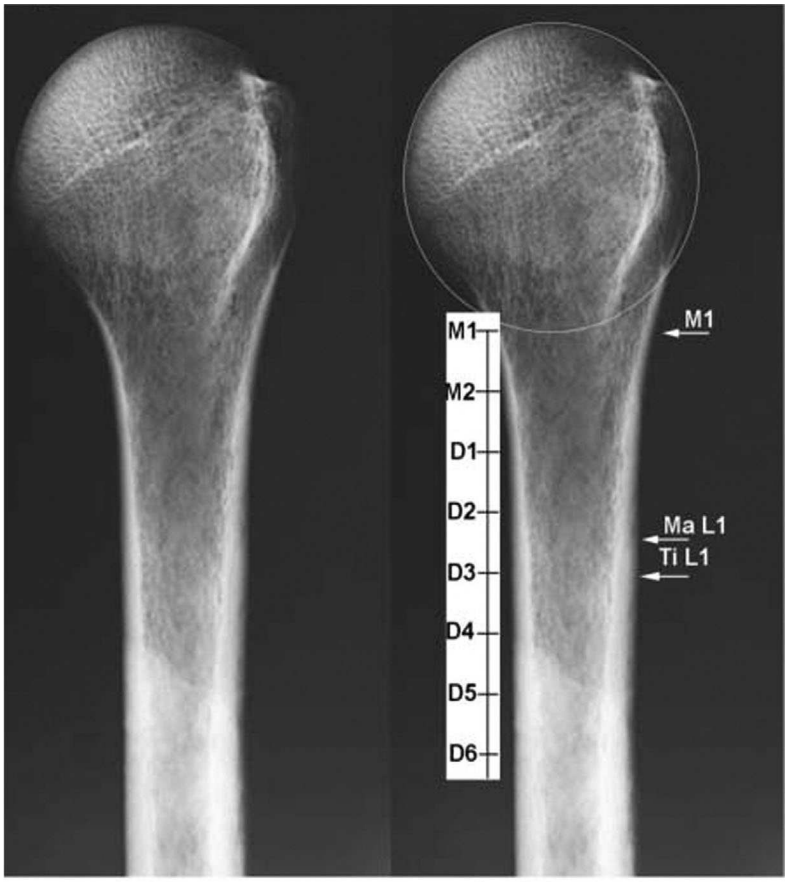


Figure 1. Two A–P radiographs from Tingart et al. [5] with original levels shown (top images). It is clear in the upper left image that the endosteal margins are not parallel. Hence the designation of Level 1 and Level 2 would be expected to be difficult, if not impossible, to make. The bottom two images show the CFM, and the more distal locations that were measured in successive 10 mm increments. The bottom two images also show the Tingart (Ti) and Mather (Ma) rectangles that were fit to each image by one of our observers after the levels shown in the original images were digitally removed. Comparisons between the two methods used in the top and bottom images are reported at the end of the Results section. (Images are reproduced with permission of The British Editorial Society of Bone and Joint Surgery.)

This new method is called the humeral head Circle-Fit Method (CFM) because it uses a circle that is manually fit (superimposed) to the articular margin of the humeral head as a way to reliably establish a proximal reference point for subsequent measurements that are made more distally in metaphyseal and diaphyseal regions (Fig. 2). We hypothesized that locations for making CI and MCCT measurements could be made more reliably using the CFM when compared to the Tingart and Mather methods. We also examined the anatomical reference method of McPherson et al.[15] to see if it is more reliable than the CFM.

METHODS

With IRB approval (no. 11,755, University of Utah), this study used digitized A–P radiographs of 33 proximal humeri from modern adults from our prior studies.[1,16] Data for the length of each humerus were not available for most specimens. The mean and standard deviation of the ages of the entire sample was 59 ± 11 years (range: 39–77 years). There were 16 males (mean 57 ± 12 years) and 17 females (mean 60 ± 10 years). The soft tissues had been manually removed from the bones prior to being radiographed. The bones were grossly normal including no evidence of arthritic changes or angular deformities. The bones ranged from robust cortices (i.e., "good quality") to osteoporotic (i.e., "poor quality") as based on morphological and densitometric analyses.[1,16] As described previously in studies that used the same bones, steps were taken to render magnification artifact negligible while using a standardized anterior–posterior projection.[17–19] This was achieved by placing the humeral head directly on the X-ray receptor platform (which is equivalent to the "cassette" in conventional non-digital imaging) with the long axis of the diaphysis aligned parallel to the platform. Each bone was externally rotated to achieve the anterior–posterior alignment shown by the arrows in Figure 3. The diaphysis



Ti = Tingart; Ma = Mather; L1 = Level 1; M1 = CFM first metaphyseal level at the lower-most (transversely tangential) edge of the circle.

Figure 2. A–P radiograph showing an unmarked proximal humerus (at left) and the same radiograph with a circle fit to the articular margin of the humeral head (at right). Shown are also the: (1) first metaphyseal location (M1, surgical neck), which is defined at the lower edge of the circle, and (2) the locations were one observer placed the upper edges (levels, L1) for the Mather (Ma) and Tingart (Ti) rectangles.

was supported with modeling clay to avoid inadvertent rotation. The beam was projected orthogonal to the platform and was focused at the surgical neck.

Five observers analyzed the digitized radiographs and made CI and MCCT measurements in accordance with: (i) method of Tingart et al.,[5] (ii) method of Tingart as modified by Mather et al.,[11] (ii) method of McPherson et al.[15] (Figs. 1 and 4,), and (iv) the new CFM (Fig. 2). For comparison we also included the two radiographs illustrated by Tingart et al.[5] (Fig. 1). The reason that they were used was to see if our observers' "levels" (obtained after placing the rectangles [described below]) on these Tingart images would be close to the levels that Tingart et al.[5] actually made and illustrated. No additional data was recorded from these two Tingart images.

The principal investigator (JGS) trained all of the observers to analyze proximal humerus radiographs. The observers included a Ph.D. graduate student in biomechanics who was studying morphology of the human humerus (and had written a Master's thesis on the topic), a medical student, two research assistants, and one board-certified orthopaedic surgeon with fellowship training in shoulder surgery (JGS). All observers had been involved in prior studies that involved examination of radiographs of the proximal humerus.[14,20] The training included several sessions where all observers: (i) observed and discussed how to locate endosteal and periosteal cortical margins in a series of radiographs of five bones spanning a broad range of bone quality, and (ii) made CI and MCCT measurements independently on these same bones and subsequently discussed and compared their results with each other and with the principal investigator (JGS). One endpoint of this training was that each observer was within 2% of the measurements made by JGS when using the CFM.

After the training phase was completed, each observer was given a compact disc (CD) with all 33 of the de-identified digitized A–P radiographs used in the present study. The order of the images on the CD was also random (i.e., they were not listed by age, sex, etc.). As described below, each observer then digitally placed the two rectangles on each radiograph, one for the Tingart method and one for the Mather method. Each observer then digitally fit a circle to the articular margin of the humeral head of each radiograph. The rectangles and circle were applied manually using the computer "mouse" and the ImageJ program (ImageJ 1.49v).[21] The circles and rectangles were fit by each observer who viewed the digitized radiographs at a magnification that was two-times actual size in a darkened room, and all measurements were corrected for magnification.[5]

The "point tool" function in ImageJ was used for fitting the circle to the articular margin of the humeral head in the CFM. This allowed the observers to manually select three points along the articular surface of the humeral head, two points at the medial and lateral edges of the articular surface and one in the middle. A circle could then be automatically created by connecting these three manually selected points. The circle size could then be readily adjusted using the computer "mouse". This method resembles the methods of

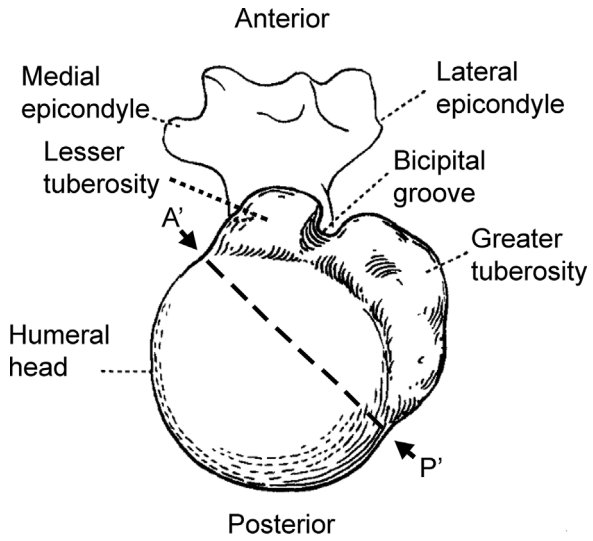


Figure 3. Drawing of a slightly oblique superior-to-inferior view of a right humerus. The X-ray beam was projected along the direction of the dotted line (A' to P'), which was achieved by externally rotating the bone. Hence after external rotation of the bone, A' is anterior and P' is posterior. The arrows therefore indicate the locations where the lesser and greater tuberosities merge with the spherical portion of the upper humerus (indicated by the arrows). The drawing is adapted from Ref.[18]

several prior studies that have considered the articular surface of the humeral head as essentially circular in standard A–P view.[17,19,22–29]

The observers used each of the following methods on each radiograph:

Tingart Method

Tingart et al.[5] state that:

> The lateral and medial cortical thickness of the proximal humeral diaphysis was measured at two different levels. Level 1 was the most proximal level of the humeral diaphysis where the endosteal borders of the lateral and medial cortices were parallel to each other. Level 2 was 20 mm distal to level 1.

In the present study, CI and MCCT measurements were made by each observer at the Tingart (Ti) levels 1 and 2 as described in Tingart et al.[5] (i.e., Ti L1 and Ti L2).

Mather Method

Mather et al.[11] state:

> The first level was the most proximal point on the humerus where the outer medial and lateral cortical borders become parallel, as previously described by Tingart et al.[5] [Note that this is an error; Tingart et al. used the endosteal (inner) cortical borders]. A perpendicular line was drawn from the medial outer cortex of the humerus to the lateral outer cortex of the humerus and measured with a digital caliper to provide the thickness of the entire bone (M1) [This is different from the "M1" used in the present study]. At the same level, the width of the intramedullary canal was obtained (M2).

Mather et al. also stated that they "denoted the parallelism of the outer proximal humerus cortex at levels 1 and 2 if the bone thickness measurements were not more than 1.0 mm different between levels." In the present study, CI and MCCT measurements were made by each observer at the Mather (Ma) levels 1 and 2 (i.e., Ma L1 and Ma L2).

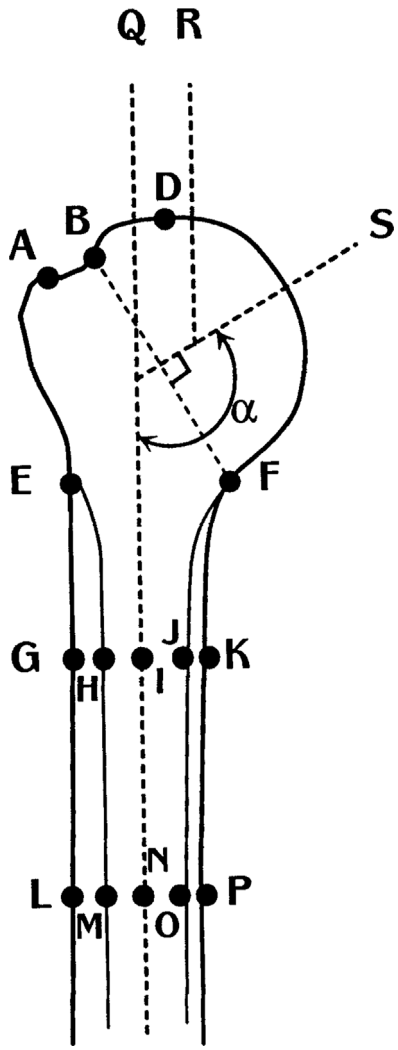


Figure 4. Drawing of A–P radiograph of proximal humerus from McPherson et al.[15] In this method, a line drawn from the points designated as E and F is an important proximal reference location for many of the subsequent measurements that they made in their extensive study of humerus morphology based on A–P and lateral radiographs. (Image is reproduced with permission of The Association of Bone and Joint Surgeons.)

McPherson Method

McPherson et al.[15] established a proximal transverse line as a reference location for subsequent measurements in their anatomical study of proximal A–P radiographs of adult

humeri (specific ages were not provided). As shown in Figure 4, they considered the two proximal reference points to be the locations where the medial and lateral proximal cortices taper to an extent that they are very thin. They illustrated this tapering as being distinct, and hence readily discernible, in the medial and lateral cortices in the proximal metaphyseal region of the humerus. These locations are shown as E and F in their drawing (Fig. 4). Because McPherson et al.[15] did not provide a rigorous definition for the E–F locations (they only illustrated them), we defined the E and F locations to be where the cortices become 1 mm thick. We then used the E–F line in a way similar to M1 of the CFM (as described below).

Circle Fit Method (CFM)

As described above, this new method employs an adjustable digitized circle that is visually best fit to the curvature of the articular surface of the humeral head on A–P radiographs (Fig. 2). The lower-most edge of the circle was defined as the proximal reference level (or "M1"; M = metaphysis = surgical neck region). All subsequent measurements that were then made distal to the M1 location were made using the ImageJ program. This included the designation of seven successive "levels" (locations) that were separated by 10 mm. This produced a second metaphyseal level (M2) and six "D" (diaphyseal) levels; the levels included M1, M2, D1, D2, D3, D4, D5, and D6. Medial and lateral cortical thickness data were also obtained and compared at each of the D levels. As shown previously, significant changes in medial or lateral cortical thickness measurements can occur when the bone is rotated $\pm 10^{\circ}$ along its long axis but not at $\pm 5^{\circ}$, which is well within the range of error that might occur when orienting the bone.[1] It took our trained observers 7–10 min to complete all measurements for one bone when using the CFM.

All Methods (Longitudinal Axis and Reliability Measurement)

For each bone (and for all measurement methods), the CFM was used to establish a longitudinal axis. The longitudinal axis was established by drawing a line that intersects the midpoint of the outer bone diameter at levels D1 and D6. Linear distances, measured in millimeters, were then made from upper-most (proximal tangent) edge of the humeral head (point D in Fig. 4) to CFM levels M1, Tingart level 1, Mather level 1, and McPherson E–F line. These linear distances were made parallel to the longitudinal axis. Intra- and inter-observer variations (referred to as "errors") in these distances were used to express reliability of each of the four reference methods.

Using ImageJ, lines were drawn perpendicular (transverse) to the longitudinal axis. These transverse lines were used to measure the outer bone diameter, the inner bone diameter (medullary canal), and the medial and lateral cortical thicknesses at: (i) Tingart levels 1 and 2, (ii) Mather levels 1 and 2, and (iii) CFM levels and D1–D6. As stated above, the E–F line of the McPherson method was considered to be similar to the M1 line (i.e., surgical neck region) of the CFM. This allowed for the E–F line to be used like the M1 line was used in the CFM. Consequently, the bone diameters and cortical thicknesses were measured with respect to the E–F line at diaphyseal locations measured similar to those shown in Figure 2.

CI and MCCT were then measured at each of the D levels designated by the CFM and McPherson method and at the L1 and L2 levels of the Tingart and Mather methods. MCCT was measured as described in Tingart et al.:[5] "the combined cortical thickness was calculated as a mean of the medial and lateral cortical thickness at the two levels and adjusted for the magnification factor." When using the CFM and the McPherson method, MCCT was calculated with respect to levels D1 and D2, and then again for D2 and D3. L1 and L2 were used for the Tingart and Mather methods.

Statistical Analysis

Statistical analysis was done using NCSS 10.[30] Inter-observer errors was analyzed for all bones ($n = 33$), males only ($n = 16$), and females only ($n = 17$). Inter-observer errors are expressed as: (1) differences in Pearson correlations between observers as described previously,[24,31] and (2) calculated differences between observers for each measurement as follows. For example, if the five observers are designated as A, B, C, D, and E, then the CFM M1 location reported with a mean and standard deviation (in mm) represents the average of following values for all five observers as follows: (A–B), (A–C), (A–D), (A–E), (B–C), (B–D), (B–E), (C–D), (C–E), and (D–E). The same approach was used for each of the other reference methods. Intra-observer error was determined by repeating all measurements for 10 bones for each observer. A one-way ANOVA design followed by Fisher's PLSD tests was then performed to determine if the variations ("errors") between the methods were significantly different.

RESULTS

The mean and standard deviation of the humeral head diameters (based on the mean of the best-fit circle drawn by the five investigators) was 49.1 ± 4.5 mm (range: 40.2–60.3 mm).

Results of CFM, Tingart, and Mather Methods

There was relatively minimal observer variation (inter and intra) in establishing the surgical neck (M1) location using the CFM. The magnitude of the intra-observer errors in establishing the M1 locations was also significantly lower ($p < 0.001$) when compared to errors incurred when the Tingart and Mather methods (Fig. 5) were each used to establish their respective L1 locations. Representative intra-observer errors are shown in Figure 6.

The relative differences in inter-observer errors among the methods were similar when comparing all bones (Fig. 5a) to males only (Fig. 5b) and females only (Fig. 5c). However, there was a tendency for greater variation in the female bones when using the Tingart and Mather methods, which might be related to greater medial versus lateral cortical asymmetry in the female bones (described below).

The relatively large errors incurred by the Tingart method significantly exceeded those of the Mather methods ($p = 0.05$). Because the L2 levels are determined by adding 20 mm in the Tingart and Mather levels the results were similar to those reported for the L1 levels (data not shown). Results with respect to the McPherson method are reported below.

There was also substantial variation with respect with which of the D levels designated in the CFM was

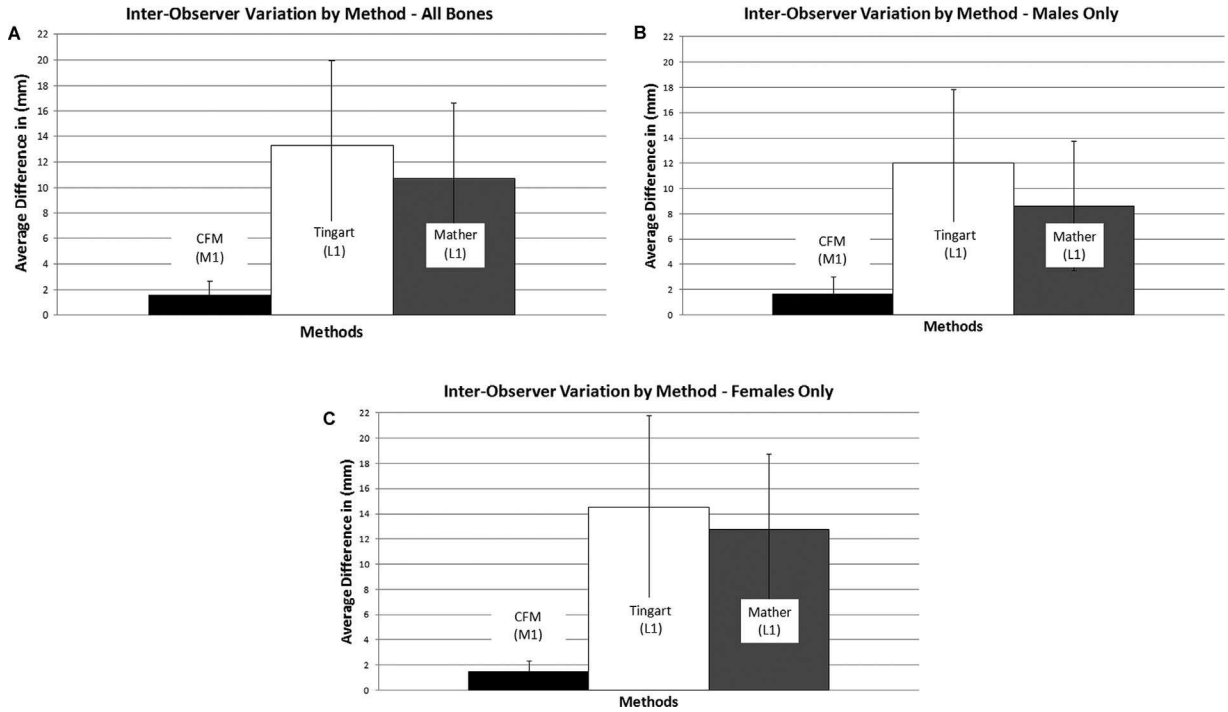


Figure 5. (A) All bones, (B) males only, and (C) females only. Inter-observer variations (averages and standard deviations) incurred when designating the M1 location using the CFM and the upper edges of the rectangles (i.e., the L1 levels) used in the Tingart and Mather methods. The broad bars represent the averages and the vertical lines are the errors, which represent the standard deviations for the differences between all five observers for each image (see Methods for details). The data are expressed in mm of the variation in the linear/longitudinal distance from the upper edge of the humeral head to each reference location. Variations for the McPherson et al. method (measured with respect to their "E–F" line) are not shown because of the absence of these landmarks in a very large majority of bones in our sample. Consequently, this method is deemed invalid as an anatomical reference method.

in closest proximity to the Tingart or Mather levels (reflecting the high variations in establishing the two levels of these two methods). For example, 33% of the time the D1 level determined using the CFM was closest to the Tingart L1. Tingart L1 was closest to D2 27% of the time and closest to D3 19% of the time. The Mather L1 level was closest to D1 44% of the time, and 26% of the time to D2 and 11% of the time to D3.

Results of the correlation analyses shown in Table 1 also support the greater reliability of the CFM. For example, the CFM showed that eight of the ten inter-observer comparisons had r values that were >0.8 (six were >0.9). By contrast, the r values never exceeded 0.8 for the Tingart and Mather methods.

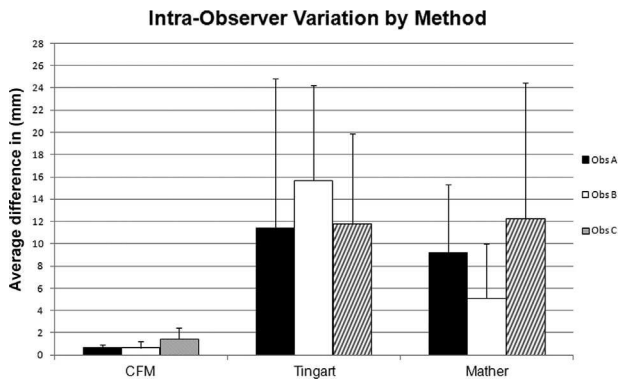


Figure 6. Representative intra-observer variations (means and standard deviations). Shown are the results of three observers (obs.) that represent the least, median, and most reliable of the five observers. See legend of Figure 4 for additional explanation of this figure.

CI data obtained using the CFM also showed less variation among the observers when the measurements were made at the D1 and D2 levels. The differences in CI among the observers are: (i) D1 $21 \pm 11\%$, D2 $21 \pm 9\%$, Mather L1 $27 \pm 11\%$, Tingart L1 $31 \pm 11\%$. Results of statistical comparisons show that CI at D1 is different from Tingart L1 ($p < 0.001$) and Mather L1 ($p = 0.002$), and at D2 is different from Tingart L1 ($p < 0.001$) and Mather L1 ($p = 0.01$). Mather and Tingart CI at their L1 are not statistically significantly different ($p = 0.12$), and CI at D1 versus D2 is also not significantly different ($p = 0.6$).

Additional results of inter-observer differences in terms of CI and MCCT are shown in Table 2. These results are shown for the entire sample of bones and not for each sex because of the non-significant relative differences in errors between the male and female bones.

Differences in cortical thickness at the D1 levels made using the CFM are show in Table 3 for all bones, males only, and females only. These results show cortical asymmetry at most locations, with the lateral cortex being thicker than the medial cortex. There is also a tendency for greater asymmetry in females compared to males. This corresponds to the greater range of error seen in females, when compared to males, when using the Tingart and Mather methods to

Table 1. Pearson Correlation Coefficients of All Five Observers When Using Each Method (CFM, Mather, and Tingart) to Designate a Reference Location*

Pearson Correlation Coefficients (r Values), Inter-Observer (Obs.) Comparisons					
	Obs. A	Obs. B	Obs. C	Obs. D	Obs. E
CFM M1					
Obs. A	1.00				
Obs. B	0.80	1.00			
Obs. C	0.85	0.96	1.00		
Obs. D	0.77	0.95	0.92	1.00	
Obs. E	0.79	0.97	0.94	0.96	1.00
Mather L1					
Obs. A	1.00				
Obs. B	0.79	1.00			
Obs. C	0.61	0.60	1.00		
Obs. D	0.65	0.58	0.79	1.00	
Obs. E	0.47	0.56	0.43[a]	0.50	1.00
Tingart L1					
Obs. A	1.00				
Obs. B	0.63	1.00			
Obs. C	0.55	0.73	1.00		
Obs. D	0.57	0.52	0.31[c]	1.00	
Obs. E	0.22[b]	0.53	0.63	0.26[b]	1.00

*This Was Done by Measuring from the Top of the Humeral Head to the Location in mm.
[a]$p = 0.01$; all other p values for the CFM and Mather method.
[b]$p > 0.14$ for these two correlations.
[c]$p = 0.08$ for this correlation.

establish the measurement locations (compare middle and right-side bars in Fig. 5b and c). By contrast the error remained consistently low when using the CFM for all bones, males only, and females only (left bars in Fig. 5a–c).

Table 2. Results of Paired Comparisons of Cortical Index (CI) and Mean Combined Cortical Thickness (MCCT) Data at Locations D1 and D2 of the CFM Method, and at Levels 1 of Tingart (Ti) and Mather (Ma) Methods. Cortical Indexes Used Here Are the Average of All Five Observers

A. Cortical Index (CI) Data	
Locations[a]	p Value
D1 vs. D2	0.04
D1 vs. Ma L1	0.04
D1 vs. Ti L1	<0.001
D2 vs. Ma L1	0.9
D2 vs. Ti L1	0.09
Ma L1 vs. Ti L1	0.08

B. MCCT (Mean Combined Cortical Thickness) Data	
Locations[a]	p Value
D1–D2 vs. D2–D3	0.2
D1–D2 vs. Ma L1–L2	0.08
D1–D2 vs. Ti L1–L2	0.03
D2–D3 vs. Ma L1–L2	0.6
D2–D3 vs. Ti L1–L2	0.4
Ma L1–L2 vs. Ti L1–L2	0.7

[a]The D levels were determined using the CFM.

Results of McPherson Method

Inter- and intra-observer errors were exceedingly high when the McPherson method was used to establish the proximal reference location (i.e., the E–F line). One reason for this is that in approximately 85% of the bones the E and F locations were relatively indistinct. In fact, the clarity with which they are depicted in Figure 4 was not seen in a very large majority of the bones that we examined. Consequently, when compared to the CFM, and even the other two methods, it was not possible to similarly employ the orientation system of McPherson and co-workers. For example, in only six of our bones (6/33 bones = 18%) was the EF line within $\pm 5^{\circ}$ of having the expected transverse orientation shown in Figure 4. For these reasons there were large errors (exceeding those that occurred with the Tingart method) among the observers when they subsequently attempted to designate the distal levels for making CI and MCCT measurements (data not shown).

Comparisons to Level 1 of Tingart Images

The differences in the L1 levels (upper edge of rectangle) between those in the original Tingart images (at top of Fig. 1) and our observers L1 levels measured using the Tingart method (at bottom of Fig. 1) are 9% greater for the left image and 3% less for the right image. These distances are with respect to the top edge of the humeral head to the levels.

Table 3. Results of Medial Versus Lateral Cortical Thickness Comparisons in All Bones, Males Only, and Females Only

Cortical Thickness Data (mm) Means and (Standard Deviations)			
All bones			
Level	Medial	Lateral	*p* Value
D1	2.6 (0.7)	2.9 (0.8)	0.09
D2	2.7 (0.7)	3.3 (0.8)	<0.001
D3	3.0 (0.8)	3.5 (1.0)	<0.001
D4	3.2 (0.8)	3.7 (1.1)	<0.001
D5	3.4 (0.9)	4.0 (1.1)	<0.001
D6	3.5 (1.1)	4.4 (1.2)	<0.001
Males only			
Level	Medial	Lateral	*p* Value
D1	2.9 (0.8)	3.0 (0.8)	0.6
D2	3 (0.8)	3.4 (0.9)	0.01
D3	3.2 (0.8)	3.7 (1.2)	0.01
D4	3.3 (0.8)	3.8 (1.3)	0.02
D5	3.5 (1.0)	4.2 (1.3)	0.01
D6	3.8 (1.1)	4.6 (1.3)	<0.001
Females only			
Level	Medial	Lateral	*p* Value
D1	2.4 (0.6)	2.7 (0.7)	0.03
D2	2.5 (0.7)	3.1 (0.7)	0.001
D3	2.9 (0.7)	3.3 (0.8)	0.001
D4	3.1 (0.9)	3.6 (0.9)	0.004
D5	3.3 (0.9)	3.8 (1.0)	0.003
D6	3.3 (1.0)	4.2 (1.1)	<0.001

Cortical asymmetry is seen at most diaphyseal (D1) locations. There is a tendency for greater asymmetry in the female bones.

DISCUSSION

Results of this study show poor reliability when using the methods of Tingart et al.[5] and Mather et al.[11] to determine locations for measuring CI or MCCT on clinical A–P radiographs of the proximal humerus. This poor reliability contrasts to the high reliability of the CFM, which was the case regardless of potential sexual dimorphism (shown in this study as medial-lateral cortical asymmetry). The poor reliability of the other two methods was expressed in terms of high variability in the reproducibility of intra- and inter-observer measurements. In addition to reducing these errors to acceptably low levels, the new humeral head CFM can also be employed using modern digital radiographic programs in the clinical setting. By contrast, we found that the E–F reference location of the McPherson et al.[15] method to be so errant that it could not be used (discussed below).

Our pilot studies have shown that the relatively high variations in CI and MCCT that occur when these measurements are made when using the Tingart and Mather methods can be clinically important. This is because these relatively high observer errors can potentially change clinical interpretation. This is show by the variations in how strongly each of these morphological parameters correlate with bone strength (i.e., ultimate fracture load) data obtained from the same bones.[14] In contrast, the minor variations that occur when CI and MCCT are made using the CFM do not similarly change the interpretation of bone strength.[14] Assuming that accuracy in CI and MCCT measurements is represented by minimal variation that they have with respect to how strongly these parameters correlate with bone strength, then we suggest that the CFM is not only more reliable but is also more accurate than the Tingart and Mather methods.

In addition to the Tingart, Mather, and the new CFM, various other methods have been described for establishing anatomical landmarks and axes for making morphological measurements of the proximal humerus.[12,15,22,24–26,32,33] McPherson et al.[15] made extensive measurements on AP and lateral radiographs of ninety-three cadaveric proximal humeri which were from "relatively elderly cadavers" (specific ages were not provided). Several of the parameters that they measured were made with respect to distances from the "humeral neck" and from the "endosteal diameter at the neck." However, the exact methods/definitions for establishing these locations are not described. But their illustration (Fig. 4) shows that they used a proximal reference location, designated as an "E location" and an "F location," where the medial and lateral cortices taper and become very thin in the proximal metaphysis. Their illustration also showed that the E and F locations are at the same transverse location of the proximal metaphysis. However, our data revealed that the variation in establishing the E and F locations at the proximal metaphysis location was even greater than the errors that were incurred when we used the Tingart and Mather methods. This is because the E and F locations at the same transverse level in accordance with McPherson et al.[15] appears to be an unrealistic simplification. In fact, in the very few instances that we found the E and F locations to be sufficiently distinct they were typically at substantially different transverse levels of the bone. Consequently, when drawing a perpendicular line from the center of the E–F line segment and extending it distally, this perpendicular line (i.e., reflecting the longitudinal axis of the bone) often deviated so greatly that it fell outside of the bone margins in some cases. By contrast, we found that when using the CFM the middle of the bone diameter at D1 and D6 can be used to reliably define and draw the longitudinal axis of each bone.

We assumed that the size of the humeral head (i.e., diameter of best-fit circle) scales proportionally with diaphyseal length. In other words, we assumed that changes in humeral head diameter scale proportionally with respect to humerus diaphysis length, which means that specific M and D levels will be at about the same percentage of total length of the humerus regardless of the size of the bone. In a study of 39 humeri from modern human humeri (ages not

reported), Roberts et al.[19] reported that the correlation coefficient between humeral head diameter and total humerus length is 0.615. We re-analyzed the data in their Figure 3 to determine the p value for the regression, which was not provided in their study. The r-value from our re-analysis of their data is 0.621 ($p < 0.001$). (This re-analysis revealed their use of 111 data points even though they reported only using 39 bones.) Nevertheless, this moderate correlation suggests that this relationship is sufficiently strong for the general purposes of our study. However in a sample of human humeri from 780 adult of various races (spanning from the Holocene to the 20th century), Auerbach and Ruff[34] report that humeral head diameter and humerus max length are poorly correlated ($r = 0.182$, $p < 0.05$). Due to conflicting data and the uncertainty of the Roberts et al. data, additional studies are needed to see if there is proportionality between the humeral head diameter and the locations where CI and MCCT are made.

Some studies provide data showing that the normal shape of the humeral articular surface is almost spherical, hence is nearly circular in the coronal plane.[22,25,26] In these studies a sphere (or circle) is used to determine the radius of curvature. But some studies note that the articular surface of the humeral head (adult bones) is slightly elliptical, with the coronal (frontal) plane being slightly larger than the sagittal plane (by ~2 mm in a large majority of cases).[12,25,32,33,35,36] Because the humeral head is only very slightly non-spherical, many investigators have considered the humeral head articular surface, as seen in A–P radiographs, as closely approximating a circle.[12,22,24–26,29] These findings support the CFM as a valid way to obtain a proximal reference point for making additional metaphyseal and diaphyseal cortical thickness and diameter measurements. In cases where the humeral head is deformed the use of the CFM might be invalid when making comparisons to bones with normal morphology.[12,27]

In a recent study,[1] we have found that the MCCT correlates more strongly with UFL at the CFM diaphyseal locations D5 and D6 when compared to D1 and D2. In this context it is important to emphasize that the Tingart and Mather proximal levels (L1) are closest to D1 and D2. This shows that the L1 reference locations obtained from using the Tingart and Mather methods are not as clinically meaningful as are the D5 and D6 locations, which are 3–4 cm farther distally (and also 1–2 cm more distal with respect to the Tingart and Mather L2 levels).

One of the limitations of the present study is that lateral radiographs were not used. It is possible that the CFM could be used on lateral radiographs. However, as noted above, there are data showing that the humeral head is less circular in lateral view when compared to a standard A–P view. Additional studies are needed to establish the use of the CFM in this context. Another limitation is that the bones examined were only from adults. We are now conducting additional studies on a much larger sample of bones (14 years to >90 years). These morphological data will also be subsequently correlated with fracture data obtained from the same bones.[1] Additional limitations in the experimental designs to obtain the ultimate fracture load data referred to above are considered in our prior study.[1] Additional studies that address sex and laterality (left vs. right humeri) in these contexts are also needed.

The results of the present study suggest that there could be novel ways to more reliably measure CI and MCCT in other long bones that have relied on measurement locations based on endocortical parallelism, diaphyseal landmarks, or percentages of bone or diaphyseal length. Various studies in different bones where consideration could be given are listed above in the Introduction section and also in others including more recent studies that have summarized some of the relevant literature.[13,37–43]

In conclusion, we found minimal inter-observer variations (mean error, 1.5 ± 1.4 mm) when the CFM was used to establish diaphyseal locations for making CI and MCCT measurements when compared to each of the Tingart and Mather methods (mean error range, 10.7 ± 5.9 to 13.3 ± 6.7 mm, respectively) ($p < 0.001$). When compared to the relatively minor variations in the CFM, the variations in the other methods are also clinically relevant because they can adversely affect the interpretation of the relationships of CI and MCCT with ultimate fracture load.

AUTHORS' CONTRIBUTIONS

All authors have read and approved the final submitted manuscript. All authors participated in the acquisition of radiographic data and writing of the manuscript.

ACKNOWLEDGMENTS

The authors thank Alex Drew and Taylor Brady for serving as observers in this study. None of the authors received monetary benefit or other sources of funds for this project.

REFERENCES

1. Skedros JG, Knight AN, Pitts TC, et al. 2016. Radiographic morphometry and densitometry predict strength of cadaveric proximal humeri more reliably than age and DXA scan density. J Orthop Res 34:331–341.
2. Nho SJ, Brophy RH, Barker JU, et al. 2007. Management of proximal humeral fractures based on current literature. J Bone Joint Surg 89:44–58.
3. Nho SJ, Brophy RH, Barker JU, et al. 2007. Innovations in the management of displaced proximal humerus fractures. J Am Acad Orthop Surg 15:12–26.
4. Spross C, Kaestle N, Benninger E, et al. 2015. Deltoid tuberosity index: a simple radiographic tool to assess local bone quality in proximal humerus fractures. Clin Orthop Relat Res 473:3038–3045.
5. Tingart MJ, Apreleva M, von Stechow D, et al. 2003. The cortical thickness of the proximal humeral diaphysis predicts bone mineral density of the proximal humerus. J Bone Joint Surg Br 85:611–617.

6. Hepp P, Theopold J, Osterhoff G, et al. 2009. Bone quality measured by the radiogrammetric parameter "cortical index" and reoperations after locking plate osteosynthesis in patients sustaining proximal humerus fractures. Arch Orthop Trauma Surg 129:1251–1259.
7. Namdari S, Voleti PB, Mehta S. 2012. Evaluation of the osteoporotic proximal humeral fracture and strategies for structural augmentation during surgical treatment. J Shoulder Elbow Surg 21:1787–1795.
8. Giannotti S, Bottai V, Dell'osso G, et al. 2012. Indices of risk assessment of fracture of the proximal humerus. Clin Cases Miner Bone Metab 9:37–39.
9. Osterhoff G, Diederichs G, Tami A, et al. 2012. Influence of trabecular microstructure and cortical index on the complexity of proximal humeral fractures. Arch Orthop Trauma Surg 132:509–515.
10. Bloom RA, Laws JW. 1970. Humeral cortical thickness as an index of osteoporosis in women. Br J Radiol 43:522–527.
11. Mather J, MacDermid JC, Faber KJ, et al. 2013. Proximal humerus cortical bone thickness correlates with bone mineral density and can clinically rule out osteoporosis. J Shoulder Elbow Surg 22:732–738.
12. DeLude JA, Bicknell RT, MacKenzie GA, et al. 2007. An anthropometric study of the bilateral anatomy of the humerus. J Shoulder Elbow Surg 16:477–483.
13. Patterson J, Rungprai C, Den Hartog T, et al. 2016. Cortical bone thickness of the distal part of the tibia predicts bone mineral density. J Bone Joint Surg Am 98:751–760.
14. Langston TD, Mears CS, Phippen CM, et al. Inter-observer variations when using popular methods to obtain cortical index and mean combined cortical thickness in proximal humerus radiographs can result in highly variable correlations with fracture strength. 62nd Annual Meeting of the Orthopaedic Research Society 2016;41: p. 1655.
15. McPherson EJ, Friedman RJ, An YH, et al. 1997. Anthropometric study of normal glenohumeral relationships. J Shoulder Elbow Surg 6:105–112.
16. Skedros JG, Pitts TC, Knight AN, et al. 2014. Reusing cadaveric humeri for fracture testing after testing simulated rotator cuff tendon repairs. Biores Open Access 3:250–254.
17. Boileau P, Walch G. 1997. The three-dimensional geometry of the proximal humerus. Implications for surgical technique and prosthetic design. J Bone Joint Surg Br 79:857–865.
18. Tillett E, Smith M, Fulcher M, et al. 1993. Anatomic determination of humeral head retroversion: the relationship of the central axis of the humeral head to the bicipital groove. J Shoulder Elbow Surg 2:255–256.
19. Roberts SN, Foley AP, Swallow HM, et al. 1991. The geometry of the humeral head and the design of prostheses. J Bone Joint Surg Br 73:647–650.
20. Phippen CM, Langston TD, Mears CS, et al. Clinical radiographic projections of the upper humerus can result in substantial errors when quantifying the deltoid tuberosity index, cortical index, and other morphological characteristics: a controlled study in cadaveric humeri. 62nd Annual Meeting of the Orthopaedic Research Society 2016;41: p. 2148.
21. Rasband W. 1997. 2016. ImageJ. Bethesda, Maryland: U.S: National Institutes of Health.
22. Robertson DD, Yuan J, Bigliani LU, et al. 2000. Three-dimensional analysis of the proximal part of the humerus: relevance to arthroplasty. J Bone Joint Surg Am 82-A: 1594–1602.
23. Soslowsky LJ, Flatow EL, Bigliani LU, et al. Articular geometry of the glenohumeral joint. Clin Orthop Relat Res 1992:181–190.
24. Yamaguchi K, Sher JS, Andersen WK, et al. 2000. Glenohumeral motion in patients with rotator cuff tears: a comparison of asymptomatic and symptomatic shoulders. J Shoulder Elbow Surg 9:6–11.
25. Pearl ML. 2005. Proximal humeral anatomy in shoulder arthroplasty: implications for prosthetic design and surgical technique. J Shoulder Elbow Surg 14:99S–104S.
26. Pearl ML, Volk AG. 1996. Coronal plane geometry of the proximal humerus relevant to prosthetic arthroplasty. J Shoulder Elbow Surg 5:320–326.
27. Youderian AR, Ricchetti ET, Drews M, et al. 2014. Determination of humeral head size in anatomic shoulder replacement for glenohumeral osteoarthritis. J Shoulder Elbow Surg 23:955–963.
28. Alolabi B, Youderian AR, Napolitano L, et al. 2014. Radiographic assessment of prosthetic humeral head size after anatomic shoulder arthroplasty. J Shoulder Elbow Surg 23:1740–1746.
29. Poppen NK, Walker PS. 1976. Normal and abnormal motion of the shoulder. J Bone Joint Surg Am 58:195–201.
30. Hintze J. 2015. NCSS 10 Statistical Software. Kaysville, Utah, USA: NCSS, LLC.
31. Kuo TY, Skedros JG, Bloebaum RD. 2003. Measurement of femoral anteversion by biplane radiography and computed tomography imaging: comparison with an anatomic reference. Invest Radiol 38:221–229.
32. Hertel R, Knothe U, Ballmer FT. 2002. Geometry of the proximal humerus and implications for prosthetic design. J Shoulder Elbow Surg 11:331–338.
33. Wirth MA, Ondrla J, Southworth C, et al. 2007. Replicating proximal humeral articular geometry with a third-generation implant: a radiographic study in cadaveric shoulders. J Shoulder Elbow Surg 16:S111–S116.
34. Auerbach BM, Ruff CB. 2006. Limb bone bilateral asymmetry: variability and commonality among modern humans. J Hum Evol 50:203–218.
35. Iannotti JP, Gabriel JP, Schneck SL, et al. 1992. The normal glenohumeral relationships. An anatomical study of one hundred and forty shoulders. J Bone Joint Surg Am 74:491–500.
36. Saka M, Yamauchi H, Hoshi K, et al. 2015. Reliability and validity in measurement of true humeral retroversion by a three-dimensional cylinder fitting method. J Shoulder Elbow Surg 24:809–813.
37. Bloom RA. 1980. A comparative estimation of the combined cortical thickness of various bone sites. Skeletal Radiol 5:167–170.
38. Dorr LD, Faugere MC, Mackel AM, et al. 1993. Structural and cellular assessment of bone quality of proximal femur. Bone 14:231–242.
39. Gruen T. 1997. A simple assessment of bone quality prior to hip arthroplasty: cortical index revisited. Acta Orthop Belg 63:20–27.
40. Baumgartner R, Heeren N, Quast D, et al. 2015. Is the cortical thickness index a valid parameter to assess bone mineral density in geriatric patients with hip fractures? Arch Orthop Trauma Surg 135:805–810.
41. Tawada K, Iguchi H, Tanaka N, et al. 2015. Is the canal flare index a reliable means of estimation of canal shape? Measurement of proximal femoral geometry by use of 3D models of the femur. J Orthop Sci 20:498–506.
42. Webber T, Patel SP, Pensak M, et al. 2015. Correlation between distal radial cortical thickness and bone mineral density. J Hand Surg Am 40:493–499.
43. Marchi D. 2010. Articular to diaphyseal proportions of human and great ape metatarsals. Am J Phys Anthropol 143:198–207.