

Original Article

Reliability of upper-limb diaphyseal mineral and soft-tissue measurements using peripheral Quantitative Computed Tomography (pQCT)

Mark A. Jenkins^{1,2,3}, Nicolas H. Hart^{2,3,4,5}, Timo Rantalainen^{2,3,4,5,6}, Paola Chivers^{2,3,4,5}, Robert U. Newton^{1,2,4}, Sophia Nimphius^{1,2,3}

¹Centre for Exercise and Sports Science Research, Edith Cowan University, Perth, Australia; ²School of Medical and Health Science, Edith Cowan University, Perth, W.A., Australia; ³Western Australian Bone Research Collaboration, Perth, W.A., Australia; ⁴Exercise Medicine Research Institute, Edith Cowan University, Perth, Australia; ⁵Institute for Health Research, The University of Notre Dame Australia, Fremantle, W.A., Australia; ⁶Gerontology Research Center, University of Jyväskylä, Jyväskylä, Finland

Abstract

Objectives: To quantify between-day reliability of upper-body diaphyseal measurements (radius, ulna, humerus) using peripheral Quantitative Computed Tomography (pQCT). **Methods:** Fourteen males (age: 25.8±2.3 years,) underwent repeat pQCT scans (one to two days apart) at mid-shaft ulna (60%), mid-shaft radius (60%) and mid-shaft humerus (50%) cross-sections of the non-dominant limb. Intraclass correlation coefficients (ICC) and coefficients of variation (CV) were determined for musculoskeletal morphology variables. **Results:** Reliability was excellent (ICC: 0.76–0.99; CV: 1.3–7.3) at all sites for bone mass, stress-strain index, endocortical and pericortical radius, endocortical volumetric bone mineral density (vBMD), muscle area, total area, non-cortical area, and cortical area. Reliability was good to excellent (ICC: 0.58–0.80; CV: 0.6–3.7) for polar vBMD and mid-cortical vBMD; fair to excellent (ICC: 0.30–0.88; CV: 0.5–8.0) for muscle density and cortical density; and fair to good (ICC: 0.25–0.60; CV: 3.4–7.6) for pericortical vBMD. Average reliability across the three sites was excellent (ICC ≥0.77; CV ≤8.0). **Conclusions:** Overall between-day reliability of pQCT was excellent for the mid-shaft ulna, radius and humerus. pQCT provides a reliable and feasible body composition and skeletal morphology assessment tool for upper limb longitudinal investigations in scientific and clinic settings.

Keywords: Bone, Ulna, Humerus, Radius, Reliability

Introduction

Peripheral Quantitative Computed Tomography (pQCT) produces a series of two-dimensional scans reconstructed to provide a three-dimensional image at specific cross-sections to determine volumetric measures of bone material, structure and strength, along with muscle and

fat morphology of the upper and lower limbs¹⁻³. The clinical utility of pQCT to assess segmental tissue composition, the effects of bone disease(s), and osteogenic adaptations due to exercise and growth is increasing², thus it is important to quantify the reliability of pQCT at various sites and across commonly assessed variables in order to determine which measures provide consistent and dependable results⁴. If pQCT results are unreliable, incorrect measurements may lead to misdiagnoses, false-positive or false-negative outcomes due to scan variability, and may compromise the ability of clinicians and researchers to accurately assess the efficacy of interventions.

Few studies have described the reliability of pQCT measures^{5,6}, principally focusing on isolated lower limb sites. Specifically, excellent reliability for cortical and trabecular bone mineral density and stress strain index (SSI) was observed when scanning the second metatarsal of cadavers⁵; and for bone mineral density in the subchondral

The authors have no conflict of interest.

Corresponding author: Dr Nicolas H. Hart, PhD, AES, CSCS, ESSAM, Senior Research Fellow, Exercise Medicine Research Institute, Edith Cowan University, 270 Joondalup Drive, JOONDALUP, Perth, Western Australia, Australia, 6027
E-mail: n.hart@ecu.edu.au

Edited by: E. Paschalis
Accepted 10 August 2018



tibia of healthy individuals⁶; with another study measuring between-day reliability of upper limb and lower limb bone sites, demonstrating high reliability with low coefficients of variation (CV)⁷. Although this single study focused on the CV's of multiple variables at several bone sites for each of the long bones in the limbs⁷, there is a need for additional measures of reliability more generally, in particular with the upper-limbs and soft-tissue. Indeed, root mean squared coefficient of variation (CV_{RMS}) values for bone measures have been reported for the upper and lower limbs⁷, however intraclass correlation coefficients (ICC) have not. Similarly, low CV_{RMS} values for various pQCT muscle measures in the lower leg⁸, and the upper-arm⁹ have also been reported, though have not yet been studied in the forearm or thigh segments to date. Similarly, there appears to be a lack of information concerning the between-day reliability of pQCT measures. Accordingly, the purpose of this investigation was to determine the between-day reliability of pQCT on commonly measured bone and muscle diaphyseal variables for the upper limb (consisting of the upper arm and forearm).

Method

Participants

Fourteen male participants (age: 25.79 ± 2.27 years, height: 179.11 ± 6.66 cm, and weight: 86.41 ± 22.84 kg), aged between 20 to 30 years volunteered and completed the current study. Males were purposefully recruited as a voluntary sub-cohort of a larger study which involved an osteogenic training intervention limited to men (i.e. minimum strength level requirement). All participants were healthy in accordance with a pre-exercise medical questionnaire to ensure no pre-existing medical conditions (i.e. current or previous fractures in measured scan sites; and movement disorders or similar), no recent nuclear medicine or radiation exposure (within the prior month), and no contraindicated medications (i.e. patients on anti-resorptive medication, such as bisphosphonates or denosumab) were present at the time of study involvement. Individuals were also excluded from participation if there was a pre-existing or recent injury (within 6 months of recruitment) in the scanned limb; if they were unable to adopt the required scan position(s) due to pain or discomfort from an existing injury to a non-scan site, if they had ever fractured or broken the scanned bones being examined. The study was approved by the University Human Research Ethics Committee (ID: JENKINS-11690) and all participants read an information sheet and provided written informed consent.

pQCT

Reliability for several variables was determined by performing test re-test pQCT scans (XCT-3000; Stratec Medizintechnik, Pforzheim, Germany), 24 or 48 hours apart within the same hour as the initial testing day to control for any potential diurnal variation. The same technician performed all scans for every session. The non-dominant

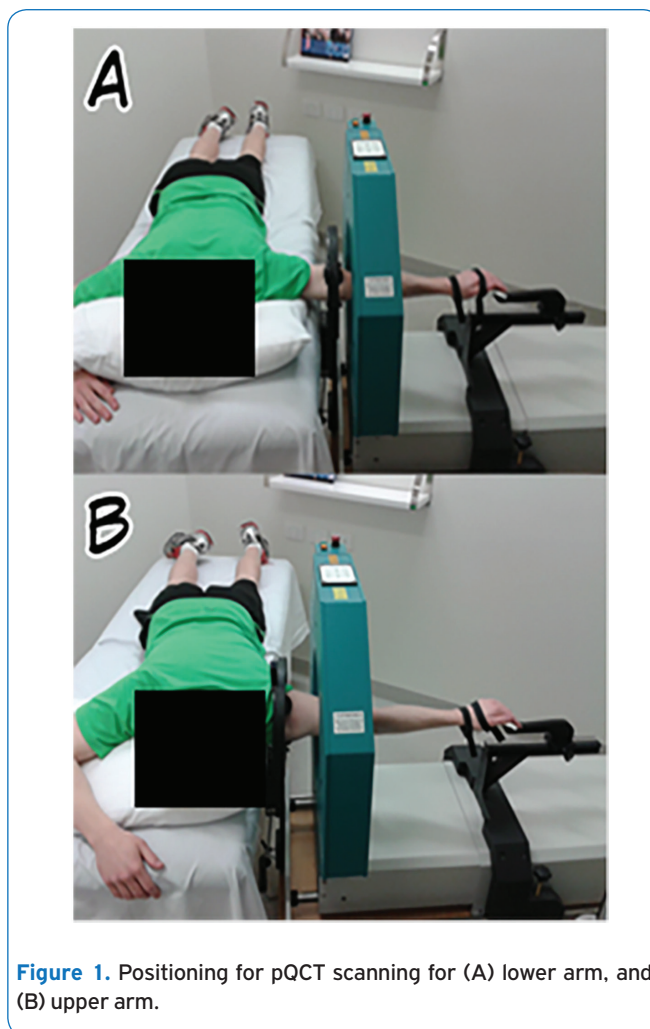


Figure 1. Positioning for pQCT scanning for (A) lower arm, and (B) upper arm.

forearm and upper arm was scanned, determined as the opposite limb of each participant's preferred writing hand. Calibration was performed daily in accordance with manufacturer specifications¹⁰. Participants were placed in a prone position on a massage table with their upper arm and/or forearm abducted 90 degrees from the torso with the elbow extended and palm face down and hand fixed inside the pQCT gantry (Figure 1). A prone position was chosen over a supine position to ensure consistency in arm orientation between upper arm and forearm scans. Specifically, when an individual has their forearm scanned while being seated, their arm is in a prone position, plus the hand fixation device was contoured to the palm of the hand in a prone position, with a strap support around the back of the hand; which is routinely used for radial scans. This attachment was thus also used for the humeral scan (observable in Figure 1). The pQCT was height adjusted for each participant to ensure the arm remained in line with the torso.

Humeral and forearm lengths were measured in triplicate (Humerus: ICC: 0.99 (0.98–1.00), CV: 0.5 (0.4–0.7); Forearm: ICC: 1.00 (0.99–1.00), CV: 0.3 (0.3–0.5)), with the average

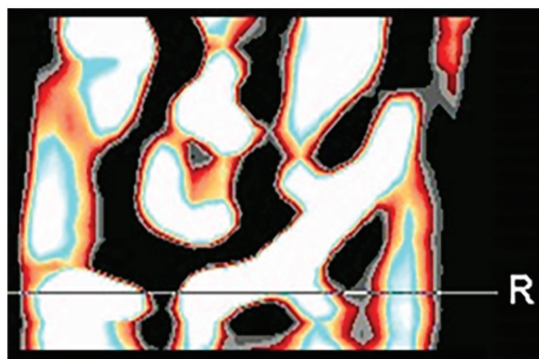


Figure 2. Scout view of the distal radioulnar joint for the forearm scan with the reference line (R) noted.

of the measures taken for each region used to establish the cross-section locations to be examined. Specifically, humeral length was defined as the lateral epicondyle at the humeroradial joint (distal end) to the acromion process at the glenohumeral joint (proximal end), with scanning occurring at the 50% humeral site. Forearm length was defined as the cortical end plate of the radius (distal end) to the lateral epicondyle at the humeroradial joint (proximal end), with scanning occurring at 60% of radial length and ulna length from the distal end. A voxel size of 0.5 mm was used for all scans. To determine the scan commencement point for the radius and ulna, a 30 mm scout view image was generated with the reference line positioned through the cortical end plate at the distal radioulnar joint (Figure 2). To determine the scan commencement point for the humerus, the halfway location of the humerus was marked on the surface of the participant's skin, with the gantry manually positioned at the marked location. An example of the scanned bones can be viewed in Figure 3. As no conspicuous anatomical landmark is visible at the distal end of humerus using the scout view owing to the olecranon process of the ulna visually obstructing the humeral endpoint, and given the added complexity of positioning a participant to access the 50% humeral slice within the gantry's moveable limits, a scout view was not used for the humerus only. Instead, the research mask was used to manually position the gantry at 50% of humeral length, which was marked on the arm from the measurement of the upper arm length.

Bone morphology was assessed using ImageJ (Version 1.48c; National Institute of Health, United States of America), BoneJ (Version 1.3.10; Imperial College London, United Kingdom) and the pQCT plug-in¹¹⁻¹³. Bone mass (mg), endocortical radius (mm), pericortical radius (mm), polar volumetric bone mineral density (vBMD; (mg/cm³), endocortical vBMD (mg/cm³) and pericortical vBMD (mg/cm³) were distributed over 36 segments (two-dimensional images rotating at ten degrees to achieve 360 degrees) and then split into 90 degree portions (0-90°; 90-180°;

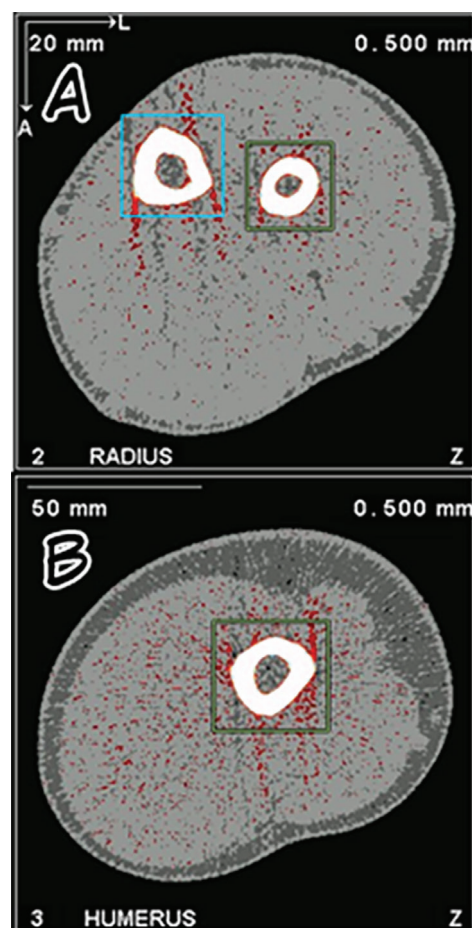


Figure 3. pQCT cross-sections (A) showing radius (green box) and ulna (blue box), and (B) Humerus.

180-270°; 270-360°) to allow for comparison of reliability between the two scanning days and in order to replicate the approach of a previous study¹⁴. The ICC and CV values for the four 90 degree portions for each variable were averaged to determine the overall reliability. Muscle density (mg/cm³), muscle area (mm²), SSI (mm³), total area (mm²), non-cortical area (mm²), cortical area (mm²), and cortical density (mm³) are unable to be broken into segmented portions, so reliability values were determined by comparing their overall values only. Additional variables (endocortical radius, pericortical radius, polar vBMD, endocortical vBMD, mid-cortical vBMD (mg/cm³), pericortical vBMD, and non-cortical area) were included which had not previously had reliability reported. Endo-, mid-, and pericortical vBMD were measured by contouring the cortex with a threshold of 220 mg/cm³, and subsequently peeling away the outer-most and inner-most layers of pixels. Thereafter, the cortical ring was evenly divided into three concentric radial rings, labelling the inner-most portion as endocortical, outer-most portion as pericortical, and the middle portion as mid-cortical, as

Table 1. Intraclass Correlation Coefficient (ICC) and Coefficient of Variance (CV) for peripheral Quantitative Computer Tomography with 90% confidence intervals for bone site variables.

	60% radius site		60% ulna site		50% humerus site	
	ICC (90% CI)	CV% (90% CI)	ICC (90% CI)	CV% (90% CI)	ICC (90% CI)	CV% (90% CI)
BM (mg)	0.96 (0.89–0.99)	2.8 (2.1–4.3)	0.99 (0.98–1.00)	2.0 (1.5–3.2)	0.99 (0.98–1.00)	2.0 (1.5–3.2)
SSI (mm ³)	0.98 (0.95–0.99)	2.7 (2.0–4.3)	0.97 (0.92–0.99)	4.3 (3.2–6.8)	0.82 (0.40–0.95)	6.6 (4.5–13.0)
EnR (mm)	0.98 (0.95–0.99)	2.8 (2.1–4.4)	0.97 (0.92–0.99)	4.8 (3.6–7.5)	0.98 (0.92–1.00)	2.5 (1.7–4.8)
PeR (mm)	0.95 (0.85–0.98)	1.6 (1.2–2.4)	0.92 (0.80–0.97)	2.3 (1.7–3.7)	0.91 (0.68–0.98)	2.4 (1.6–4.5)
PvBMD (mg/cm ³)	<i>0.63 (0.29–0.14)</i>	<i>0.6 (0.5–0.9)</i>	<i>0.58 (0.20–0.82)</i>	3.7 (2.7–5.7)	0.77 (0.34–0.94)	1.6 (1.1–3.2)
EnvBMD (mg/cm ³)	0.80 (0.54–0.92)	3.2 (2.4–5.0)	0.76 (0.47–0.91)	6.6 (4.9–10.4)	0.86 (0.54–0.97)	2.9 (2.0–5.7)
MivBMD (mg/cm ³)	0.80 (0.56–0.92)	1.6 (1.2–2.5)	0.69 (0.33–0.88)	<i>1.7 (1.3–2.7)</i>	<i>0.66 (0.15–0.90)</i>	<i>1.0 (0.7–1.8)</i>
PevBMD (mg/cm ³)	0.25* (-0.24–0.64)	4.3 (3.2–6.7)	0.34* (-0.09–0.68)	7.6 (5.6–12.0)	<i>0.60 (-0.01–0.88)</i>	<i>3.4 (2.3–6.5)</i>
MD (mg/cm ³)	0.30* (-0.20–0.68)	0.7 (0.6–1.2)	0.30* (-0.20–0.68)	0.7 (0.6–1.2)	0.88 (0.58–0.97)	0.5 (0.3–0.9)
MuA (mm ²)	0.99 (0.98–1.00)	1.3 (1.0–2.1)	0.99 (0.98–1.00)	1.3 (1.0–2.1)	0.99 (0.98–1.00)	2.0 (1.4–3.9)
ToA (mm ²)	0.98 (0.94–0.99)	1.7 (1.2–2.6)	0.96 (0.89–0.99)	2.8 (2.1–4.4)	0.98 (0.90–0.90)	1.9 (1.3–3.7)
MeA (mm ²)	0.99 (0.97–1.00)	4.0 (2.9–6.2)	0.98 (0.95–0.99)	7.3 (5.4–11.6)	0.99 (0.97–1.00)	3.1 (2.2–6.1)
CoA (mm ²)	0.97 (0.91–0.99)	2.0 (1.5–3.1)	0.94 (0.84–0.98)	3.7 (2.7–5.7)	0.97 (0.87–0.99)	2.7 (1.9–5.3)
CoD (mm ³)	0.31* (-0.20–0.68)	8.0 (5.9–12.7)	<i>0.58 (0.15–0.83)</i>	3.4 (2.5–5.3)	<i>0.74 (0.21–0.93)</i>	<i>2.3 (1.6–4.5)</i>
Average	0.77 (0.58–0.85)	2.6 (2.0–4.1)	0.78 (0.58–0.91)	3.7 (2.8–5.9)	0.87 (0.61–0.96)	2.5 (1.7–4.8)

BM: bone mass; SSI: stress-strain index; EnR: endocortical radius; PeR: pericortical radius; PvBMD: polar volumetric bone mineral density; EnvBMD: endocortical volumetric bone mineral density; MivBMD: Mid-cortical volumetric bone mineral density; PevBMD: pericortical volumetric bone mineral density; MD: muscle density; MuA: muscle area; ToA: total area; MeA: non-cortical area; CoA: cortical area; CoD: cortical density. *Italicised values represent fair to good ICC, * represents poor ICC values whereas all other values were considered excellent (ICC > 0.75).*

seen in Rantalainen et al., 2011¹⁵. Bone area was contoured at 220 mm² and these pixels were used for structure-related measurements (CoA, ToA, SSI). Subsequently volumetric density was defined as all pixels within the bone envelope with density higher than or equal to 220 mg/cm³. For density distribution analyses the dense cortex was contoured with 220 mg/cm³ and the outermost layers of pixels were subsequently peeled away and disregarded. These contouring selections are the norm in the field. The lower threshold used for area-related measures ensures accurate area measures¹⁶, and a higher threshold is used for volumetric density measures to minimise influence of the partial volume effect (defined as pixels that are partially filled with soft-tissues, and/or partially with hard-tissues).

Statistical analysis

Between-day reliability of pQCT was assessed by using a reliability analysis spreadsheet (Analysis of Reliability Spreadsheet, Version 2012, SportScience, New Zealand) to determine the ICC and the CV with 90% confidence and upper and lower confidence limits^{17,18}. Data was assessed for outliers based on visual identification of acceptable scans from previous studies^{19,20}. Two outliers were identified and removed for the radius and ulna analyses, while seven outliers were identified and removed for the humerus analysis. The most likely cause for the outliers is attributed to motion artefact. Statistical analyses were conducted with SPSS (Version 19.0.0; IBM SPSS Statistics, USA) and a

reliability analysis program by Hopkins (2000)^{17,18}. An ICC value greater than 0.75 was regarded as having excellent reliability, an ICC value between 0.4 and 0.75 inclusive was regarded as having fair to good reliability, and an ICC value below 0.4 was regarded as having poor reliability²¹. Further, a CV less than 10% was considered acceptable²².

Results

ICC and CV for variables assessed by pQCT at the three different sites (60% mid-shaft ulna; 60% mid-shaft radius; 50% mid-shaft humerus) across fourteen different variables (bone mass; SSI; endocortical radius; pericortical radius; polar vBMD; endocortical vBMD; mid-cortical vBMD pericortical vBMD; muscle area; muscle density; total area; non-cortical area; cortical area; cortical density) are summarised in Table 1. ICC and CV results for 90 degree segments are summarised in Table 2.

The overall ICC value for the radius was excellent with 0.77, with a CV value of 2.6%. For the radius, all ICC values for bone mass, SSI, endocortical radius, pericortical radius, total area, non-cortical area, and cortical area, were excellent and ranged from 0.95 to 0.99, with CV values ranging from 0.9% to 4.0%. Endocortical vBMD and mid-cortical vBMD both had excellent ICC values of 0.80, with CV values of 3.2% and 1.6% respectively. Polar vBMD had a fair to good ICC value of 0.63, and a CV value of 0.6%. Pericortical vBMD and cortical density had a fair to good ICC values ranging from 0.25 to

Table 2. Intraclass Correlation Coefficient (ICC) and Coefficient of Variance (CV) for peripheral Quantitative Computer Tomography with 90% confidence intervals for 90 degree segments.

	60% radius site		60% ulna site		50% humerus site	
	ICC (90% CI)	CV% (90% CI)	ICC (90% CI)	CV% (90% CI)	ICC (90% CI)	CV% (90% CI)
BM (mg) 0-90°	0.96 (0.90-0.99)	2.2 (1.6-3.4)	0.99 (0.97-1.00)	2.2 (1.7-3.5)	0.97 (0.86-0.99)	3.4 (2.4-6.7)
90-180°	0.96 (0.90-0.99)	2.5 (1.9-3.9)	0.99 (0.98-1.00)	1.5 (1.1-2.3)	0.94 (0.78-0.99)	4.0 (2.7-7.7)
180-270°	0.95 (0.87-0.98)	2.7 (2.0-4.3)	0.99 (0.98-1.00)	1.8 (1.4-2.9)	0.98 (0.92-1.00)	2.9 (2.0-5.6)
270-360°	0.96 (0.89-0.98)	3.6 (2.7-5.7)	0.99 (0.97-1.00)	2.5 (1.8-3.9)	0.99 (0.94-1.00)	2.4 (1.6-4.6)
EnR (mm) 0-90°	0.97 (0.92-0.99)	3.2 (2.4-5.0)	0.97 (0.92-0.99)	4.4 (3.3-6.9)	0.98 (0.93-1.00)	2.3 (1.6-4.5)
90-180°	0.98 (0.95-0.99)	2.5 (1.9-3.9)	0.98 (0.94-0.99)	4.5 (3.4-7.1)	0.99 (0.94-1.00)	2.2 (1.5-4.2)
180-270°	0.98 (0.96-0.99)	2.5 (1.8-3.8)	0.97 (0.92-0.99)	4.9 (3.6-7.6)	0.99 (0.96-1.00)	1.7 (1.2-3.3)
270-360°	0.99 (0.96-1.00)	3.1 (2.3-4.9)	0.96 (0.88-0.98)	5.4 (4.0-8.4)	0.97 (0.86-0.99)	3.6 (2.5-7.1)
PeR (mm) 0-90°	0.94 (0.83-0.98)	1.7 (1.3-2.6)	0.98 (0.95-0.99)	1.5 (1.1-2.4)	0.96 (0.83-0.99)	1.7 (1.2-3.3)
90-180°	0.96 (0.89-0.99)	1.4 (1.0-2.1)	0.78 (0.48-0.92)	3.9 (2.9-6.1)	0.82 (0.41-0.96)	3.3 (2.2-6.3)
180-270°	0.89 (0.71-0.96)	2.1 (1.5-3.2)	0.97 (0.91-0.99)	2.0 (1.5-3.1)	0.90 (0.64-0.98)	2.4 (1.6-4.6)
270-360°	0.99 (0.97-1.00)	1.1 (0.8-1.7)	0.95 (0.86-0.98)	1.9 (1.4-3.0)	0.96 (0.82-0.99)	2.0 (1.3-3.8)
PvBMD (mg/cm³) 0-90°	0.94 (0.83-0.98)	1.8 (1.3-2.8)	0.88 (0.70-0.96)	2.4 (1.8-3.7)	0.86 (0.51-0.97)	0.5 (0.4-1.0)
90-180°	0.82 (0.56-0.93)	2.8 (2.1-4.3)	0.23* (-0.28-0.64)	5.2 (3.8-8.1)	0.68 (0.10-0.91)	1.9 (1.3-3.7)
180-270°	0.48 (0.01-0.78)	2.8 (2.1-4.4)	0.55 (0.10-0.81)	3.1 (2.3-4.8)	0.94 (0.77-0.99)	1.3 (0.9-2.4)
270-360°	0.28* (-0.23-0.67)	1.8 (1.3-2.8)	0.65 (0.26-0.86)	3.9 (2.9-6.2)	0.61 (-0.02-0.89)	2.8 (1.9-5.5)
EnvBMD (mg/cm³) 0-90°	0.94 (0.85-0.98)	4.2 (3.1-6.5)	0.88 (0.70-0.96)	8.1 (6.0-12.8)	0.89 (0.59-0.97)	1.1 (0.8-2.2)
90-180°	0.89 (0.72-0.96)	3.3 (2.4-5.1)	0.51 (0.05-0.79)	6.5 (4.8-10.3)	0.73 (0.19-0.93)	3.8 (2.6-7.3)
180-270°	0.68 (0.31-0.87)	3.0 (2.2-4.7)	0.81 (0.55-0.93)	3.3 (2.5-5.2)	0.98 (0.92-1.00)	2.4 (1.7-4.7)
270-360°	0.67 (0.29-0.87)	2.3 (1.7-3.6)	0.83 (0.59-0.94)	8.3 (6.2-13.2)	0.84 (0.44-0.96)	4.4 (3.0-8.7)
MivBMD (mg/cm³) 0-90°	0.98 (0.94-0.99)	0.9 (0.7-1.5)	0.81 (0.54-0.93)	1.5 (1.1-2.4)	0.37* (-0.33-0.80)	1.3 (0.9-2.6)
90-180°	0.91 (0.76-0.97)	2.2 (1.6-3.4)	0.74 (0.40-0.90)	1.2 (0.9-1.8)	0.94 (0.78-0.99)	0.6 (0.4-1.2)
180-270°	0.77 (0.47-0.91)	1.4 (1.1-2.3)	0.51 (0.04-0.79)	1.8 (1.3-2.8)	0.56 (-0.10-0.87)	0.9 (0.6-1.7)
270-360°	0.53 (0.08-0.81)	1.9 (1.4-2.9)	0.70 (0.34-0.88)	2.3 (1.7-3.6)	0.75 (0.25-0.94)	1.0 (0.7-1.8)
PevBMD (mg/cm³) 0-90°	0.54 (0.08-0.81)	3.6 (2.7-5.6)	0.83 (0.59-0.94)	2.0 (1.5-3.1)	0.44 (-0.25-0.83)	2.8 (1.9-5.4)
90-180°	0.09* (-0.40-0.54)	4.0 (3.0-6.3)	-0.02* (-0.49-0.46)	11.3 (8.4-18.1)	0.82 (0.39-0.95)	2.9 (2.0-5.6)
180-270°	0.31* (-0.19-0.69)	5.9 (4.4-9.2)	0.44 (-0.05-0.76)	8.8 (6.5-13.9)	0.58 (-0.06-0.88)	3.5 (2.4-6.7)
270-360°	0.06* (-0.43-0.52)	3.6 (2.7-5.7)	0.09* (-0.40-0.54)	8.1 (6.0-12.9)	0.56 (-0.10-0.87)	4.2 (2.9-8.2)
Average	0.77 (0.55-0.90)	2.6 (2.0-4.1)	0.75 (0.51-0.89)	4.1 (3.0-6.4)	0.82 (0.50-0.95)	2.4 (1.7-4.7)

BM: bone mass; EnR: endocortical radius; PeR: pericortical radius; PvBMD: polar volumetric bone mineral density; EnvBMD: endocortical volumetric bone mineral density; MivBMD: Mid-cortical volumetric bone mineral density; PevBMD: pericortical volumetric bone mineral density. Italicised values represent fair to good ICC, an * represents poor ICC values whereas all other values were considered excellent (ICC > 0.75).

0.31, with CV values ranging from 4.3% to 8.0%.

The overall ICC value for the ulna was excellent with 0.78, and a CV value of 3.7%. For the ulna, all ICC values for bone mass, SSI, endocortical radius, pericortical radius, total area, non-cortical area, and cortical area were excellent and ranged from 0.92 to 0.99, with CV values ranging from 2.0% to 7.3%. Endocortical vBMD had an excellent ICC value of 0.76, and a CV value of 6.6%. Polar vBMD, mid-cortical vBMD, and cortical density had a fair to good ICC values ranging from 0.58 to 0.69, with CV values ranging from 1.7% to 3.7%.

Pericortical vBMD had a poor ICC value of 0.34, and a CV value of 7.6%.

The overall ICC value for the humerus was excellent with 0.87, and a CV value of 2.5%. For the humerus, all ICC values for bone mass, SSI, endocortical radius, pericortical radius, polar vBMD, endocortical vBMD, total area, non-cortical area, and cortical area were all excellent and ranged from 0.77 to 0.99, with CV values ranging from 0.5% to 6.6%. Mid-cortical vBMD, pericortical vBMD, and cortical density had fair to good ICC values ranging from 0.60 to 0.74, with

CV values ranging from 1.0% to 3.4%.

For soft-tissues, the muscle area for the forearm had an excellent ICC value of 0.99, and a CV value of 1.3%. The muscle density for the forearm had a poor ICC value of 0.30, and a CV value of 0.7%. The muscle area for the upper arm had an excellent ICC value of 0.99, and a CV value of 2.0%. The muscle density for the upper arm had an excellent ICC value of 0.88, and a CV value of 0.5%.

For 90 degree segments, bone mass, endocortical radius, and pericortical radius had excellent ICC values for all segments (≥ 0.78), with low CV values ($\leq 5.4\%$). Endocortical vBMD had ICC values ranging from fair to good, to excellent ($0.51 \leq 0.98$), with low CV values ($\leq 8.3\%$). Polar vBMD, mid-cortical vBMD, and pericortical vBMD had ICC values ranging from poor to excellent ($-0.02 \leq 0.98$), with all CV values low ($\leq 8.8\%$) except for one location for the pericortical vBMD (11.3%).

Discussion

The purpose of this investigation was to determine the between-day reliability of pQCT on commonly measured diaphyseal bone and muscle variables of upper-limb (upper-arm and forearm) cross-sections. We demonstrate between-day reliability of the pQCT for bone measures to be excellent for a majority of reported variables at all upper limb sites, however several variables at various sites only provided fair to good reliability with some variables providing poor reliability. The between-day reliability for upper limb muscle measures were excellent for all muscle area sites, and poor to excellent for all muscle density sites. This study reports, for the first time, reliability data for numerous humeral measures, as well as radial and ulnar diaphyseal bone material and density distribution (endo- and pericortical radii, polar vBMD, endo-, mid- and pericortical vBMD) and forearm muscle measures. These reliability results support the utility for pQCT to provide correct, accurate diagnostic measurements of bone and muscle assessments to be made by clinicians and researchers, with subsequent interventions able to confidently demonstrate efficacy. Although SSI reliability has not been previously reported in the forearm bones, the excellent reliability results are similar to a study by Chaplais et al., 2014 which focused on the second metatarsal, also reporting excellent ICC values (≥ 0.99) and very low CV values ($\leq 8.9\%$), suggesting broadly that SSI measures in the upper limbs are reliable⁵. The excellent reliability results are also similar to a study by Weatherholt et al. (2015), which had very low CV values (1.4%), also suggesting that SSI measures in the upper limbs are reliable⁹.

Area measurements of endocortical radius and pericortical radius had excellent ICC values for all bones (≥ 0.91) and very low CV values for all bones ($\leq 4.8\%$). This is in contrast to the volumetric bone measurements of polar vBMD, endocortical vBMD, mid-cortical vBMD, and pericortical vBMD which demonstrated good reliability with low CV values ($\leq 7.6\%$), though had ICC values ranging from poor (0.25) to excellent

(0.86) across sites. Area measurements of muscle were also excellent for all upper limb sites (ICC=0.99), while poor to excellent for muscle density across sites (ICC=0.30-0.88), with high reliability demonstrated in the upper arm versus low reliability in the forearm. This is similar to the bone variables, such that area measurements exhibit high ICC values and low CV values while density measurements exhibited moderate ICC values and low CV values. A possible reason for the excellent reliability observed in area measures versus reduced reliability in density measures could be due to slight patient positioning discrepancies between visits with regards to manual positioning of the cross-section examined or parallax error due to subtle variations of the limb perpendicular to the gantry. Inherently, measures of area may be less sensitive to movement artefact or subtle positional differences between days and scans relative to density measures. Positively, no current studies, to the authors' knowledge, have measured ICC values for muscle variables, which strengthens the current study, particularly as CV's measure the standard deviation as a percent of the mean, and is suitable for within-subject variation; though an ICC takes into account both within and between-subject variation, thus appropriately strengthens the value of reliability analyses^{23,24}.

CV reliability results for upper limb bone mass, cortical area and cortical density are similar to the multiple reliability findings previously reported by Sievänen et al., 1998. Indeed, bone mass results at all three bones for both studies reported very low CV values ($\leq 4.9\%$)⁷, with similarly low CV values between studies for cortical area and cortical density values ($\leq 8.0\%$)⁷. Cortical area and cortical density for this study (Radius $\leq 8.0\%$, Ulna $\leq 3.7\%$; Humerus $\leq 2.7\%$) are comparable to Sievänen et al, 1998⁷ (Radius $\leq 6.5\%$; Ulna $\leq 2.1\%$; Humerus $\leq 4.6\%$). While the CV results for cortical area and cortical density may slightly vary between the studies, all CV values are very low and considered excellent. Furthermore, muscle area CV values for the upper-limb sites of this study ($\leq 2.0\%$) are very low and similar to the muscle area CV values in the calf region reported by Wong et. al, 2017 ($\leq 2.9\%$), muscle density values in the forearm by Frank-Wilson et al, 2015 ($\leq 5.3\%$), and muscle density values in the upper arm by Weatherholt et al, 2015 (1.9%)^{8,9,25}. The CV values are also similar for muscle density values between our study ($\leq 0.7\%$) and the same studies by Wong et. al., 2017 ($\leq 4.1\%$), Frank-Wilson et al, 2015 ($\leq 2.6\%$), and Weatherholt et al, 2015 (0.7%)^{8,9,25}. The CV values for the upper arm are also similar for total area between our study (1.9%) and a study by Weatherholt et al, 2015 (1.0%)⁹. There are possible reasons for differences in CV between muscle density and muscle area (which are two separate measurements in the first instance). Muscle area is cross-sectional area (outer circumference) of the muscle; whereas muscle density factors in the amount of muscle within the cross-sectional area defined, inclusive of intramuscular fat and other factors. In addition, CV is calculated as SD/Mean. If the scale does not start from zero (as is the case with volumetric density), a CV does not provide a reasonable description of variability. This

is why we also report ICC, which does not suffer from this scale-related anomaly. CV is reported because that is what is routinely reported in the bone field to represent reliability.

Increased reliability at the 60% mid-shaft radius site and the 60% distal ulna site compared to the 50% humerus site may be due to the commencement point for the scans of the forearm being selected by a scout view of the pQCT, while the 50% humerus site was located manually and positioned manually. While the manual selection of the 50% humerus site did result in an excellent average ICC value, it is possible that determining the scan location manually may be less precise than it would be by determining the scan location using scout view. This is because errors in the manual measurements of the limbs may lead to errors in the scan locations²⁶. While this may lead to errors with or without a scout view, previous reference lines from scout views may be used for subsequent scans and this will allow for more precise scans with the scout views; subsequent manual measurements do not have previous reference lines to use²⁶. Another possible reason for the lower average ICC values for the 50% humerus site is due to the clamping and fixation of the arm during scanning. For forearm scans, a clamp can be used to prevent movement²⁷. For the humerus scans, a clamp was not used due to the 50% humerus site being unable to reach the scan location of the pQCT whilst ensuring participant comfort. This may have resulted in slight movement artefact or subtle arm relocation which would have resulted in lower ICC value between days and scans. However, the methodological aspects to securing limbs for scanning was not evaluated in our study and certainly warrants further investigation, particularly with respect to improving reliability and minimising motion artefact. While small motion artefact may not have a large effect on the results, a large amount of motion artefact can affect the results negatively²⁶; and these may have differential effects on area variables relative to density variables. Lastly, an additional limitation and potential reason for the varying ICC values could be due to the positional expertise and measurement of the technician performing the pQCT scans.

Conclusion

The overall reliability of pQCT is excellent for the three scanned upper limb long bones as reported by both ICC values and low CV values. Humerus 50% scans are more reliable than radius and ulna 60%, and the humeral reliability could be improved by creating new limb clamping methods to prevent arm movement during the scanning process. Area measurements for bone and some muscle variables were more reliable than volumetric measurements across upper-limb sites.

Acknowledgements

The authors would like to thank the Exercise Medicine Research Institute for the use of their pQCT. MJ was supported by an Australian Government Research Training Program Scholarship. NHH is supported by the Cancer Council of Western Australia Research Fellowship.

References

1. Hart NH, Nimphius S, Rantalainen T, Ireland A, Siafarikas A, Newton RU. Mechanical basis of bone strength: influence of bone material, bone structure and muscle action. *J Musculoskel Neuronal Interact* 2017;17(3):1-26.
2. Lee DC, Gilsanz V, Wren TAL. Limitations of peripheral quantitative computed tomography metaphyseal bone density measurements. *J Clin Endocrinol Metab* 2007; 92(11):4248-53.
3. Engelke K, Adams JE, Armbrrecht G, Augat P, Bogado CE, Bouxsein ML, Felsenberg D, Ito M, Prevrhal S, Hans DB, Lewiecki EM. Clinical use of quantitative computed tomography and peripheral quantitative computed tomography in the management of osteoporosis in adults: The 2007 ISCD official positions. *J Clin Densitom* 2008;11(1):123-62.
4. Sullivan GM. A primer on the validity of assessment instruments. *J Grad Med Educ* 2011;3(2):119-20.
5. Chaplais E, Greene D, Hood A, Telfer S, du Toit V, Singh-Grewal D, Burns J, Rome K, Schiferl DJ, Hendry GJ. Reproducibility of a peripheral quantitative computed tomography scan protocol to measure the material properties of the second metatarsal. *BMC Musculoskelet Disord* 2014;15:242-7.
6. Wrigley T V, Creaby M W, Bennell K L. Peripheral quantitative computed tomography (pQCT) of tibial subchondral bone in knee osteoarthritis. *J Biomech* 2007;40(Suppl 2):S591.
7. Sievänen H, Koskue V, Rauho A, Kannus P, Heinonen A, Vuori I. Peripheral Quantitative Computed Tomography in Human Long Bones: Evaluation of *In Vitro* and *In Vitro* Precision. *J Bone Miner Res* 1998;13(5):871-82.
8. Wong AKO, Hummel K, Moore C, Beattie KA, Shaker S, Craven BC, Adachi JD, Papaioannou A, Giangregorio, L. Improving reliability of pQCT-derived muscle area and density measures using a watershed algorithm for muscle and fat segmentation. *J Clin Densitom* 2015; 18(1):93-101.
9. Weatherholt AM, Avin KG, Hurd AL, Cox JL, Marberry ST, Santoni BG, Warden SJ. Peripheral quantitative computed tomography (pQCT) predicts humeral diaphysis torsional mechanical properties with good short-term precision. *J Clin Densitom* 2015; 18(4):551-9.
10. XCT 3000 manual software version 6.20. Pforzheim: Stratec Medizintechnik GmbH; 2009.
11. National Institutes of Health, ImageJ 2016, <http://imagej.nih.gov/ij/>
12. Rantalainen T, Nikander R, Daly RM, Heinonen A, Sievänen H. Exercise loading and cortical bone distribution at the tibial shaft. *Bone* 2011;48(4):786-91.
13. Doube M, Kłosowski MM, et al. "BoneJ: free and extensible bone image analysis in ImageJ." *Bone* 2010; 47(6):1076-9.
14. Macdonald HM, Cooper DML, McKay HA. Anterior-

- posterior bending strength at the tibial shaft increases with physical activity in boys: Evidence for non-uniform geometric adaptation. *Osteoporos Int* 2009;20(1):61-70.
15. Rantalainen T, Nikander R, Heinonen A, Daly RM, Sievänen H. An open source approach for regional cortical bone mineral density analysis. *J Musculoskel Neuronal Interact* 2011;11(3):243-8.
 16. Rittweger J, Michaelis I, Giehl M, Wüsecke P, Felsenberg, D. Adjusting for the partial volume effect in cortical bone analyses of pQCT images. *Journal of Musculoskeletal and Neuronal Interactions* 2004;4(4):436-441.
 17. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000;30(1):1-15.
 18. Sportscience, Precision of measurement 2011, www.newstats.org/precision.html
 19. Blew RM, Lee VR, Farr JN, Schiferl DJ, & Going SB. Standardizing evaluation of pQCT image quality in the presence of subject movement: qualitative versus quantitative assessment. *Calcif Tissue Int* 2014; 94(2):202-11.
 20. Rantalainen T, Chivers P, Beck BR, Robertson S, Hart NH, Nimphius S, et al. Please don't move-evaluating motion artifact from peripheral quantitative computed tomography scans using textural features. *J Clin Densitom* 2018;21(2):260-268.
 21. Fleiss JL. The design and analysis of clinical experiments. New York: John Wiley & Sons; 1986.
 22. Pedersen J, Liu J. Child mortality estimation: Appropriate time periods for child mortality estimates from full birth histories. *PLoS Med* 2012;9(8):1-13.
 23. Sportscience, Standard Deviation 2000, <http://www.sportsci.org/resource/stats/stdev.html#cv>
 24. Sportscience, Calculations For Reliability, 2009, <http://www.sportsci.org/resource/stats/relycalc.html>
 25. Frank-Wilson AW, Johnston JD, Olszynski WP, Kontulainen SA. Measurement of muscle and fat in postmenopausal women: precision of previously reported pQCT imaging methods. *Bone* 2015;75(June):49-54.
 26. Shields RK, Dudley-Javoroski S, Boaldin KM, Corey TA, Fog DB, Ruen JM. Peripheral quantitative computed tomography: measurement sensitivity in persons with and without spinal cord injury. *Arch Phys Med Rehabil* 2006;87(10):1376-81.
 27. Szabo KA, Webber CE, Adachi JD, Tozer R, Gordon C, Papaioannou A. Cortical and trabecular bone at the radius and tibia in postmenopausal breast cancer patients: A Peripheral Quantitative Computed Tomography (pQCT) study. *Bone* 2011;48(2):218-24.