Short Communication

Reliability of 2D:4D measurements using a direct method suitable for clinical settings

Laura Costas [a,b,c,*], Eleni Papadopoulou [d,e,f,g], Beatriz Perez-Gomez [h,i], Xavier Basagaña [d,e,i], Francesc Alarcon [a,i], Delphine Casabonne [a,i], Yolanda Benavente [a,i], Gemma Castaño-Vinyals [d,e,i], Manolis Kogevinas [d,e,g,i], Silvia de Sanjosé [a,b,i]

[a] Unit of Infections and Cancer, Cancer Epidemiology Research Programme, IDIBELL, Catalan Institute of Oncology, Barcelona, Spain
[b] Department of Medicine, University of Barcelona, Barcelona, Spain
[c] Department of Epidemiology, Biostatistics and Occupational Health, Faculty of Medicine, McGill University, Montreal, Canada
[d] Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain
[e] Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain
[f] Department of Experimental and Health Sciences, Pompeu Fabra University, Barcelona, Spain
[g] National School of Public Health, Athens, Greece
[h] Área de Epidemiología Ambiental y Cáncer, CNE, ISCIII, Madrid, Spain
[i] CIBER Epidemiologia y Salud Pública (CIBERESP), Madrid, Spain

ARTICLE INFO

ABSTRACT

The second to fourth digit ratio (2D:4D) indicates androgen exposure during early development. Numerous methods are used in the assessment of 2D:4D, however, some are not reliable enough and others are difficult to perform in large epidemiological studies. We assessed the reliability of 2D:4D using a direct method with a caliper, and compared it with a computer-assisted analysis on scanned images. Both methods were moderately correlated, but the scan method produced slightly lower 2D:4D. The main source of variation was differences between subjects (real variation). Reliability was higher among men and among younger participants. All reliability coefficients were higher than 0.8 when three repeated measurements were averaged. Our results suggest that reliability is influenced by participants' characteristics. Digit ratios determined directly with calipers are reliable when repeated measurements are averaged.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The ratio of the lengths of the 2nd and 4th hand digits (2D:4D) has been suggested to be a proxy indicator of hormone exposure during pre-natal development. Higher in utero testosterone levels result in lower 2D:4D (Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004; Manning, Scutt, Wilson, & Lewis-Jones, 1998), possibly mediated by a different balance of hormone receptors between digits (Zheng & Cohn, 2011).

2D:4D has been postulated as a potential indicator of fetal hormonal exposures leading to adult diseases, as suggested by the Barker hypothesis (Barker, 2004). Low 2D:4D may be related to a delayed menarche (Manning & Fink, 2011), to a decreased breast cancer risk (Muller et al., in press), and to an increased prostate cancer risk (Muller et al., 2011; Rahman et al., 2011). They may also influence eating disorders (Quinton, Smith, & Joiner, 2011), alcohol dependency (Kornhuber et al., 2011), and social risk-taking behaviors (Stenstrom, Saad, Nepomuceno, & Mendenhall, 2011).

However, these measurements could be subject to intra- and inter-observer variability. The technique used to assess 2D:4D may affect reliability by introducing random and/or systematic error. Validation studies of 2D:4D measures are generally carried out among young people, and whether age affects reliability is unknown. Our aim was to assess the reliability of 2D:4D measurements made by experienced examiners in a sample of young and middle-aged adults. We compared measurements using a direct method with a digital caliper with those using a computer-assisted analysis on scanned images of the same subjects.

2. Material and methods

2.1. Study subjects and data collection

Fifty workers (24 men and 26 women) from the Catalan Institute of Oncology in Barcelona (Spain) were recruited for this vali-

dation study. Age and sex of participants were recorded. Participants did not have any anomaly on the fingers. Between three and five experienced interviewers determined three times the length of the 2nd and 4th finger on both hands of each participant with a digital caliper, directly on the surface of the palms. Each examiner performed 600 direct measurements per volunteer. Examiners had a previous training and they followed a protocol for fingers measurement. They asked participants to remove any ring from their hands and they measured lengths from the middle point of the proximal crease to the fingertip. The screen of the caliper was covered to blind the procedure, and the measurement was recorded by another researcher. Each participant's hands were also scanned using a Hewlett Packard scanjet 5400c scanner (300 dpi resolution), and lengths of 2nd and 4th fingers were measured three times using the ImageJ Program (Rasband, 1997).

2.2. Statistical analyses

We used absolute-agreement intraclass correlation coefficients (ICC) to assess the correlation between methods. Mixed-effects models were fitted to decompose the total variance into variance components for lengths of 2nd and 4th digit, and their corresponding 2D:4D. Variance components were used to estimate within-examiner reliability coefficients. *P*-values for the comparison of reliability coefficients between groups were obtained using permutation tests (Davison & Hinkley, 1997). Confidence intervals for reliability coefficients were obtained using the bootstrap method (Davison & Hinkley, 1997). Reliability coefficients for averages of repeated measurements were computed through the Spearman–Brown equation (Lord & Novick, 1968). All analyses were stratified by sex.

3. Results

Mean age of participants was 35 (range 21–52). Direct and scan methods showed a moderate to high agreement, with ICCs ranging from 0.71 to 0.96 in men and from 0.64 to 0.95 in women (Table 1), although scan method was more precise (ICC range = 0.64–0.96 for caliper and ICC range = 0.88–0.97 for scan method). Ratios were slightly lower with the scan method than with digital caliper (median difference = 0.007; range −0.122 to 0.099). Women had

slightly higher 2D:4D than men, but this was only observed in the right hand (Table 1).

The between-subjects variance (real variance) was the main component of the total variance of the measurements (Table 2). Reliability was high for both digit lengths and 2D:4D, although within- and between-examiner variance components were high for 2D:4D, and they reached up to 37%. Reliability of 2D:4D was higher for men than for women (Table 2, P = 0.093), and slightly higher for age below 35 than for 35 or older (0.70 vs. 0.52, P = 0.17). Nevertheless, reliability increased with repeated measurements, and all reliability coefficients were higher than 0.8 when three repeated measurements were averaged.

4. Discussion

2D:4D ratios may indicate hormone exposure during human early development and they have been associated to numerous hormone-related conditions. However, these measures could be subject to intra- and inter-observer variability. Our results showed that direct measurements made by experienced examiners were reliable: all reliability coefficients were higher than 0.8 when repeated measurements were averaged. A recent report by Zheng and Cohn (2011) provided insights on the biology of this sexually dimorphic feature, and revealed that the 2D:4D ratio in mice is mediated by androgen receptor (AR) and estrogen receptor α (ER-α) activity. The 4th digit has more ARs and ERs-α than the 2nd digit, and the balance of this signaling modulates the levels of skeletogenic gene expression during a narrow window of digit development. Consistent with our data, Honekopp and Watson observed in a meta-analysis that 2D:4D presented larger sex differences in the right hand compared to the left hand, and concluded that right hand ratios are possibly a better indicator of prenatal hormone exposure (Hönekopp & Watson, 2010). Interestingly, Zheng and Cohn (2011) observed that only the right paw showed significant differences between males and females during the embryonic period, suggesting that this is produced by unknown biological factors intra-utero rather than by handedness later in life.

These traits could help in the identification of fetal origins of adult diseases, as they reflect the endocrine signaling during the prenatal period. They have been associated to conditions such as

Table 1
Anthropometric characteristics by method and level of agreement between methods.

	Caliper		Scan		ICC (between methods)
	Mean	SD	Mean	SD	
Men (n = 24)					
Left hand					
2D (mm)	73.3	4.5	74.0	4.6	0.96
4D (mm)	74.9	4.7	76.3	4.8	0.91
2D:4D	0.979	0.038	0.971	0.037	0.83
Right hand					
2D (mm)	73.6	4.7	74.2	4.7	0.96
4D (mm)	75.3	4.4	76.4	4.4	0.91
2D:4D	0.978	0.038	0.971	0.027	0.71
Women (n = 26)					
Left hand					
2D (mm)	66.3	4.3	67.1	4.0	0.94
4D (mm)	67.8	4.4	68.9	4.4	0.92
2D:4D	0.978	0.033	0.974	0.026	0.64
Right hand					
2D (mm)	67.1	4.2	67.7	4.0	0.95
4D (mm)	67.7	4.3	69.1	4.4	0.91
2D:4D	0.991	0.035	0.980	0.028	0.74

mm = Millimeters.
ICC = Intraclass correlation coefficient.