Validity and reliability of dual-energy X-ray absorptiometry for the assessment of abdominal adiposity

Scott G. Glickman,1,2 Charles S. Marn,3 Mark A. Supiano,1,2 and Donald R. Dengel4,5

1Department of Internal Medicine, Division of Geriatric Medicine, University of Michigan; 2The Geriatric Research, Education, and Clinical Center and 3Department of Radiology, Ann Arbor Veterans Affairs Medical Center, Ann Arbor, Michigan 48105; 4School of Kinesiology, University of Minnesota, Minneapolis, 55455; and 5Minneapolis Veterans Affairs Medical Center, Minneapolis, Minnesota 55417

Submitted 18 November 2003; accepted in final form 1 April 2004

Glickman, Scott G., Charles S. Marn, Mark A. Supiano, and Donald R. Dengel. Validity and reliability of dual-energy X-ray absorptiometry for the assessment of abdominal adiposity. J Appl Physiol 97: 509–514, 2004. First published April 9, 2004; 10.1152/japplphysiol.01234.2003.—A number of methods exist for the estimation of abdominal obesity, ranging from waist-to-hip ratio to computed tomography (CT). Although dual-energy X-ray absorptiometry (DXA) was originally used to measure bone density and total body composition, recent improvements in software allow it to determine abdominal fat mass. Sixty-five men and women aged 18–72 yr participated in a series of studies to examine the validity and reliability of the DXA to accurately measure abdominal fat. Total body fat and abdominal regional fat were measured by DXA using a Lunar DPX-IQ. Multislice CT scans were performed between L1 and L4 vertebral bodies (region of interest) using a Picker PQ5000 CT scanner, and volumetric analyses were carried out on a Voxel Q workstation. Both abdominal total tissue mass (P = 0.02) and abdominal fat mass (P < 0.0001) in the L1–L4 region of interest were significantly lower as measured by DXA compared with multislice CT. However, Bland-Altman analysis demonstrated good concordance between DXA and CT for abdominal total tissue mass (i.e., limits of agreement = −1.56–2.54 kg) and fat mass (i.e., limits of agreement = −0.40–1.94 kg). DXA also showed excellent reliability among three different operators to determine total, fat, and lean body mass in the L1–L4 region of interest (intraclass correlations, R = 0.94, 0.97, and 0.89, respectively). In conclusion, the DXA L1–L4 region of interest compared with CT proved to be both reliable and accurate method to determine abdominal obesity.

A NUMBER OF STUDIES HAVE DEMONSTRATED that the accumulation of adipose tissue in the abdominal region is independently associated with diabetes, stroke, and coronary heart disease (8, 11, 14, 17). Despite the important role of abdominal obesity in numerous diseases, practical measurements of abdominal adipose tissue are not readily available. At present, the most accurate in vivo method of measuring abdominal adipose tissue is computed tomography (CT). Although this method represents a technological advance and is used as the reference standard, its application for body composition assessment in routine clinical practice and body composition research is limited because of expense, access to the scanner, and exposure to significant quantity of ionizing radiation.

Dual-energy X-ray absorptiometry (DXA) provides a reliable estimate of whole body composition. This technique is quick and accurate and exposes subjects to minimal amounts of ionizing radiation (10). DXA allows for separation of the body into regions of interest including the abdominal, often defined as the fat mass located between lumbar vertebral bodies L1 and L4 (7, 25, 26). Although the use of DXA to determine the adiposity in the abdominal region has gained popularity, the validity and reliability of the DXA to measure this region accurately have yet to be determined. Therefore, the purpose of the present study was to determine the validity and reliability of DXA to measure abdominal soft tissue composition, defined as the vertebral L1–L4 operator-defined region of interest.

MATERIALS AND METHODS

Experimental Design

A series of studies was conducted to determine both the validity and reliability of DXA for determination of abdominal composition. Subjects were recruited through newspapers and flyers in the community. Study participants were free of disease, not pregnant or using medication affecting body composition, and without metal implants. Subjects did not eat, drink, urinate, defecate, or exercise between tests. In all studies, subjects were clothed in either T-shirt and athletic shorts or a standard hospital gown.

The methods and procedures used in this investigation were approved by the Human Subjects’ Committee at the Department of Veterans Affairs Ann Arbor Healthcare System. All participants provided informed consent before being tested.

Study 1: Validity of DXA to determine abdominal composition. To examine the validity of the DXA L1–L4 region of interest as an accurate measure of abdominal fat, 27 subjects (15 men and 12 women) had their abdominal fat determined by both multislice CT and DXA on the same day. Subjects were randomized to either multislice CT or DXA, with the other test immediately following the first. Before the DXA or CT scan, skinfold thickness and circumferences in the abdomen region were taken by a single technician.

Study 2: Reliability of DXA to measure abdominal composition. To examine the reliability of the DXA to determine abdominal composition, two separate experiments were carried out. Study 2 examined the ability of the DXA to measure changes in body composition at the L1–L4 region of interest. Twenty-eight subjects (8 men, 20 women) were scanned before and after packets of porcine lard of uniform thickness (~2.5 cm) were placed over the subject’s L1–L4 region of interest. The amount of fat placed over the L1–L4 region of interest was calculated to approximate a 10% (0.5–1.0 kg) increase in fat in this region. The DXA scans were performed consecutively on the same day while the subject repositioned, two separate experiments were carried out. Study 2 examined the ability of the DXA to measure changes in body composition at the L1–L4 region of interest. Twenty-eight subjects (8 men, 20 women) were scanned before and after packets of porcine lard of uniform thickness (~2.5 cm) were placed over the subject’s L1–L4 region of interest. The amount of fat placed over the L1–L4 region of interest was calculated to approximate a 10% (0.5–1.0 kg) increase in fat in this region. The DXA scans were performed consecutively on the same day while the subject repositioned.
mained in the same position on the DXA table. The fat content of the porcine lard (91.84%) was determined by chemical fat extraction (methanol-chloroform extraction) (12).

**Study 3: Interrater reliability of DXA to measure abdominal composition.** Three technicians received verbal, written, and hands-on instruction outlining the analysis procedure and the location of bony landmarks visible on the digital image. Each technician manually determined the DXA L1–L4 region of interest on 43 scans.

**Anthropometry**

Body weight was measured to the nearest 0.1 kg via a medical beam scale. Height was measured to the nearest 0.5 cm using a stadiometer. Triplicate circumference measurements of the natural waist (narrowest part of the torso as seen from the anterior aspect) and umbilicus were taken by use of a Gulick hand-woven tape (1). Skinfold thicknesses were measured three times at the suprailiac, suprailium, and abdomen with a Harpenden skinfold caliper after standardization procedures (19), except for the abdomen skinfold where a vertical rather than horizontal fold was measured. The average value of the three trials was used as the criterion value.

**Multislice CT Measurements**

Multislice CT measurements were performed with a Picker PQ5000 CT scanner and analyzed on a Voxel Q workstation (Picker International, Highland Heights, OH) at 135 kVp, 100-mA exposure. Subjects were examined in the supine position with arms outstretched overhead to decrease beam hardening or streak artifact. The scanning region of interest was from the superior aspect of the first lumbar vertebral body (T12/L1 intervertebral disc) to the inferior aspect of the fourth lumbar vertebral body (L4/L5 intervertebral disc), assessed on a pilot image, and intended to correspond to the DXA L1–L4 region of interest. All spiral scans were performed with a pitch factor equal to 1.0, creating a helix of adjacent horizontal cross-sectional slices including the entire region L1–L4 with no between-slice gaps. The number of slices ranged between 16 and 19.

Evaluation of the attenuation histogram established the mean attenuation (Hounsfield units [HU]) for adipose tissue as −130 HU (range −190 to −30 HU) and −20 HU for soft tissue (range −190 to +100) (10, 11, 13). A fat tissue-highlighting technique was used to determine the subcutaneous and intra-abdominal adipose tissue areas. CT scan analysis determined fat mass via calculation of a fat volume based on sequential spiral transverse scans through the abdomen that enables separation of subcutaneous from visceral fat. With the use of a volume correction of 0.900 g/cm³ for the density of fat (13), estimates of total abdominal tissue mass and abdominal fat mass were determined.

**DXA**

Subjects were scanned by use of a whole body DXA system (model DPX-IQ, Lunar Radiation, Madison, WI; software version 4.5e) set at medium speed and medium collimation ratio. Subjects lay supine on the DXA table with arms adequately separated from the trunk and were instructed to remain still throughout the scanning procedure. After analysis of the whole body scan, a quadrilateral box was manually drawn around the L1–L4 region of interest (abdomen) bounded inferiorly by the horizontal line identifying L4/L5 vertebral space and superiorly by the horizontal line identifying the T12/L1 vertebral space. Scans were displayed with an adjustment of the gray scale, so that all of the soft tissue in the designated area was included.

**Statistical Analysis**

Univariate regression analysis was used to describe the relationship between total abdominal mass and abdominal fat mass quantified by multislice CT and DXA. Mean differences between estimates of total abdominal mass and abdominal fat mass measured by multislice CT and DXA were tested for statistical significance by paired Student’s t-test. These differences were then related to the mean of the two estimates for each variable as described by Bland and Altman (2).

To determine the sensitivity of the DXA to alterations in abdominal fat mass, simple univariate regression analysis was used to describe the relationship between total abdominal mass and abdominal fat mass determined by DXA with and without added fat. Mean differences between DXA scans with and without added fat were tested for statistical significance by paired Student’s t-tests. Total abdominal mass and abdominal fat mass were related to the mean of the two estimates as described by Bland and Altman (2).

Reliability by different DXA operators to measure regional adiposity was evaluated via a one-way ANOVA with post hoc analysis. In addition, intraclass correlations were calculated for total, fat, and lean body mass in the L1–L4 region of interest.

Stepwise multiple linear regression analyses using measures of abdominal adiposity, including DXA and anthropometry, were performed to determine the best predictors of CT-measured visceral adipose tissue. Data were analyzed with SAS and StatView 5.0 (SAS Institute, Cary, NC) and are reported as means ± SE. Significance was set at the 0.05 level for all tests.

**RESULTS**

The men and women participating in this study ranged in age from 18 to 72 yr. By design, the population studied in the present study represented a range of body compositions (8.0–58.0% body fat).

**Study 1: Validity of DXA to Determine Abdominal Composition**

Total abdominal tissue mass for the L1–L4 region of interest as measured by DXA (7.07 ± 1.96 kg) was significantly (P = 0.02) lower than that measured by multislice CT (7.48 ± 1.87 kg). There was strong correlation between multislice CT-measured total abdominal tissue mass and DXA-measured total abdominal tissue mass (r = 0.858, P < 0.0001) (Fig. 1A). The slope and intercept values of the regression lines describing the relationship between these two methods were not significantly different from 1 and 0, respectively. Bland-Altman analysis for total abdominal tissue mass is presented in Fig. 1B and represents moderate agreement (95% limits of agreement: −1.56–2.54 kg) between the two testing modalities.

Similarly, abdominal fat mass for the L1–L4 region of interest as measured by DXA (2.22 ± 1.63 kg) was significantly (P < 0.0001) lower than that measured by multislice CT (2.99 ± 1.99 kg). There was a strong correlation between CT-measured abdominal fat mass and DXA-measured abdominal fat mass (r = 0.967, P < 0.0001) (Fig. 2A). The slope and intercept values of the regression lines describing the relationship between these two methods were also not significantly different from 1 and 0, respectively. There was moderate agreement (95% limits of agreement: −0.41–1.94 kg) between the two methods as represented by the Bland-Altman plot (Fig. 2B).

**Study 2: Reliability of DXA to Measure Abdominal Composition**

DXA scans of the L1–L4 region were made with and without fat packets. As expected, there were significant differences between the DXA estimates for total tissue mass in the L1–L4 region with and without added fat (7.38 ± 3.92 vs.
6.58 ± 3.69 kg, \( P < 0.0001 \)). However, when the DXA estimate of the L1–L4 region with the added fat was corrected by subtracting the known fat packet mass, there was no significant difference between the two estimates for total tissue mass (6.63 ± 2.71 vs. 6.58 ± 3.69 kg, \( P = 0.405 \)). Similarly there were also significant differences between the DXA estimates for fat mass in the L1–L4 region with and without added fat (2.81 ± 0.25 vs. 2.26 ± 0.23 kg, \( P < 0.0001 \)). However, when the DXA of the L1–L4 region with added fat was corrected for the fat packet mass, the L1–L4 fat mass was significantly less than that measured in the original scan (2.12 ± 0.24 vs. 2.26 ± 0.23 kg, \( P < 0.0001 \)), although the difference between the mean values was very small.

**Study 3: Rater Reliability of DXA to Measure Abdominal Composition**

A total of 43 DXA scans was analyzed by each of three operators (Table 1). There were no significant differences for total mass (\( P = 0.76 \)), fat mass (\( P = 0.98 \)), or lean body mass (\( P = 0.54 \)) in the L1–L4 region of interest box when determined by any of the three operators. Intraclass correlations calculated from the entire pool of 43 scans revealed excellent reliability among the three operators for total mass (\( R = 0.94, P > 0.0001 \)), fat mass (\( R = 0.97, P > 0.0001 \)), and lean body mass (\( R = 0.89, P > 0.0001 \)).

### Table 1. Measures of body composition in the DXA L1–L4 region of interest box as determined by 3 separate operators

<table>
<thead>
<tr>
<th>Operator 1</th>
<th>Operator 2</th>
<th>Operator 3</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total mass, kg</td>
<td>6.89±2.58</td>
<td>7.18±3.25</td>
<td>6.73±2.66</td>
</tr>
<tr>
<td>Fat mass, kg</td>
<td>2.05±1.44</td>
<td>2.11±1.60</td>
<td>2.05±1.61</td>
</tr>
<tr>
<td>Lean body mass, kg</td>
<td>4.84±1.50</td>
<td>5.07±2.00</td>
<td>4.67±1.40</td>
</tr>
<tr>
<td>Bone mass, kg</td>
<td>0.95±0.29</td>
<td>1.04±0.37</td>
<td>0.92±0.31</td>
</tr>
</tbody>
</table>

Values are means ± SD. DXA, dual-energy X-ray absorptiometry.
Prediction Equations

Stepwise regression equations with the intercept passing through zero were generated by involving combinations of waist circumference, suprailiac, suprailium, and abdomen skinfolds and DXA measurements. Fat mass in the L1–L4 region of interest accounted for 80% of the variance in CT-derived visceral adipose tissue \( (r = 0.894, P < 0.0001) \). When the circumference of the natural waist was added to the model, 84% of the variance in CT derived visceral adipose tissue was accounted for \( (r = 0.915, P < 0.0001) \). None of the three skinfold measurements significantly improved the model. The final equation to predict visceral adipose tissue was

\[
\text{Visceral adipose tissue (g)} = \text{DXA L1-L4 fat mass} \times 0.31 + \text{natural waist} \times 7.03
\]

DISCUSSION

The salient findings of the present study are that 1) compared with CT, DXA provides a valid method to determine abdominal adiposity; 2) DXA sensitively detects changes in abdominal fat in the L1–L4 region; and 3) determination of the L1–L4 region of interest using DXA is highly reproducible and independent of operator error. To our knowledge, this is the first study to determine the validity and reliability of DXA to determine abdominal adiposity in the operator-determined L1–L4 region of interest.

Study 1: Validity of DXA to Determine Abdominal Composition

Currently, CT is recognized as the criterion standard for determination of abdominal adiposity. Although other studies have compared DXA’s ability to measure abdominal fat vs. CT, most of these studies have used single-slice CT scans (24, 26). Use of cross-sectional area at a single level assumes that the area of abdominal fat in the single cross section is representative of the entire abdominal region. Recently, Greenfield et al. (9) examined the validity of a single-slice CT for determination of abdominal fat and reported significant intra-subject variability in premenopausal women. In the present study, we used spiral CT to volumetrically determine abdominal fat content by accounting for the entire quantity of tissue within the measured region. Both univariate regression and Bland-Altman analyses indicated good agreement between the two methods for determining total abdominal tissue mass and abdominal fat mass, although the DXA values for each were significantly less. In the present study, CT estimates of abdominal fat mass exceeded DXA by \( \sim 26\% \). This is similar to the 20% difference between CT and DXA estimates of abdominal fat mass and DXA reported by Svendsen et al. (25), who also observed a significant correlation between CT and DXA measured abdominal fat mass and DXA measured abdominal fat mass. Even though there were differences in abdominal fat mass between the two methods, abdominal fat mass was highly correlated between DXA and CT, which is similar to results by Jensen et al. (13). Although the two methods (i.e., CT and DXA) may be highly correlated, this does not mean that the two methods agree (2). Therefore, Bland-Altman (2) analysis of the difference between the two methods was performed and demonstrated good agreement between the two methods for determining both total abdominal mass and abdominal fat mass.

Systematic differences in radiation physics between DXA and CT scanners potentially add a degree of measurement error (18). One technical problem encountered on CT measurement of fat areas is beam hardening and scatter radiation caused by bone tissue, which may lead to problems in the estimate of fat mass (27). In addition, another source of potential error is that the CT scanner assumes that there exists a linear change between adjacent cross-sectional areas (29). DXA also has technical limitations. In those individuals with high bone mineral density, fat mass may be underestimated because bone mineral content relative to lean tissue is assumed to be constant (27).

Study 2: Reliability of DXA to Measure Abdominal Composition

For DXA to be used in longitudinal or intervention studies, it must also be sensitive to changes in the abdominal region. In the present study, we examined the ability of DXA to determine changes in abdominal fat by placing packets of porcine fat over the abdomen. DXA accurately accounted for the total mass of the added porcine fat packet that was placed over the abdomen. However, DXA only accounted for 78% of the total fat of the porcine fat packet that was added. This is improved from previous reports in which DXA detected only 52 and 55% of the additional fat placed over the abdominal region (20, 23). One reason for the difference between these studies may be explained by the different DXA scanners and software that were used. Snead et al. (23) used a Hologic QDR/W DXA scanner instead of the Lunar DPX-IQ scanner tested in the present study. Earlier reports claim that these two DXA scanners produce different body composition results (28). Another possible explanation for differences between the present study and those of Snead et al. may be that Snead et al. assumed that the amount of porcine fat contained in the fat packets was 100% lipid material. However, chemical analysis of the lard used in the present study indicated its contents at \( \sim 92\% \) fat. In the present study, all calculations of added fat were done using the value determined by chemical analysis instead of 100%. Milliken et al. (20) used an earlier version of the DXA scanner (Lunar DPX-L scanner) and software (software version 1.3y) than used in the present study and reported that the Lunar DPX-L scanner underestimated the lard packet when placed on the abdomen \( (90\% \text{ actual vs. } 52\% \text{ measured}) \). When two lard packets were placed one a top of the other, there was an increase in the underestimation of the added fat \( (90\% \text{ actual vs. } 47\% \text{ measured}) \).

Study 3: Rater Reliability of DXA to Measure Abdominal Composition

Measurement of abdominal obesity using DXA requires a technician to manually draw a box around this region. Because this is an area of potential human error, we sought to evaluate the ability of different technicians to manually define this region of interest. To our knowledge this is the first study examining human error in DXA assessments of regional adiposity. The DXA estimates of total mass, fat mass, lean body mass, or bone in the L1–L4 region of interest were not significantly different among the three operators. The high
degree of agreement between the three operators reported in this study may be due to use of bone landmarks when defining the region of interest box. These results indicate that the L1–L4 region of interest box is highly reproducible and independent of operator error. It should be noted that another potential source of variability in the reproducibility of the DXA L1–L4 region of interest box may be due to positioning of subjects on the table for the scan. In the present study, we did not reposition subjects between duplicate scans and cannot comment on this potential source of error.

Prediction Equations

Although other studies have provided regression equations incorporating combinations of anthropometric, total, and regional body composition measures to derive visceral adipose tissue, these previous equations only predict the visceral adipose tissue area or the visceral adipose tissue-to-subcutaneous adipose tissue ratio (5, 15, 26), which, although correlated, does not equate with visceral adipose tissue mass. Those equations (13, 15) that predict visceral adipose tissue mass were developed from single-slice CT scans vs. the method used in the present study.

Accumulation of adipose tissue in the abdominal region, independent of total adiposity, has been associated with insulin resistance, Type 2 diabetes, and cardiovascular disease (3, 4, 21). Therefore, safe and noninvasive techniques to accurately determine abdominal adiposity are of clinical significance in the early detection of individuals at risk for metabolic and cardiovascular diseases. Results of the present study indicate that DXA measures of total abdominal fat mass agree extremely well with total adipose tissue mass measured by a volumetric CT scanning technique. The DXA L1–L4 region of interest technique is both reliable and reproducible for assessment of abdominal adipose tissue; however, its relationship to metabolic and cardiovascular disease risk factors is unknown. Future studies are needed to determine the relationship between the DXA L1–L4 region of interest measure of abdominal fat and metabolic and cardiovascular disease.

There are systematic differences between CT and DXA scanners that appear to be related to the underlying principles of the two techniques. It should be noted that even though DXA was found to be reliable and reproducible in the estimation of abdominal adipose tissue, it does not allow the visual distinction between visceral and subcutaneous fat tissue. However, combining fat mass in the L1–L4 region of interest and the circumference of the natural waist accounted for 84% of the variance in CT-derived visceral adipose tissue, providing researchers with a method to estimate visceral adipose tissue using the DXA. Appreciation of the strengths and weaknesses of each method can facilitate selecting the appropriate body composition techniques for a given study.

REFERENCES


GRANTS

This research was supported by National Institutes of Health Research Scientist Development Award in Aging KO1 AG-0072301 (D. R. Dengel) and National Institute on Aging K24 AG-00924 (M. A. Supiano), Department of Veterans Affairs Medical Research Service (M. A. Supiano), and Geriatric Research, Education and Clinical Center (D. R. Dengel and M. A. Supiano) at Ann Arbor, University of Michigan, Claude D. Pepper Older Americans Independence Center (AG-08808), University of Michigan Clinical Research Center (RR-00042).

J Appl Physiol • VOL. 97 • AUGUST 2004 • www.jap.org


