Review Article

Body-composition assessment via air-displacement plethysmography in adults and children: a review[1–3]

David A Fields, Michael I Goran, and Megan A McCrory

ABSTRACT Laboratory-based body-composition techniques include hydrostatic weighing (HW), dual-energy X-ray absorptiometry (DXA), measurement of total body water (TBW) by isotope dilution, measurement of total body potassium, and multicompartment models. Although these reference methods are used routinely, each has inherent practical limitations. Whole-body air-displacement plethysmography is a new practical alternative to these more traditional body-composition methods. We reviewed the principal findings from studies published between December 1995 and August 2001 that compared the BOD POD method (Life Measurement, Inc, Concord, CA) with reference methods and summarized factors contributing to the different study findings. The average of the study means indicates that the BOD POD and HW agree within 1% body fat (BF) for adults and children, whereas the BOD POD and DXA agree within 1% BF for adults and 2% BF for children. Few studies have compared the BOD POD with multicompartment models; those that have suggest a similar average underestimation of ≈2–3% BF by both the BOD POD and HW. Individual variations between 2-compartment models compared with DXA and 4-compartment models are partly attributable to deviations from the assumed chemical composition of the body. Wide variations among study means, −4.0% to 1.9% BF for BOD POD − HW and −3.0% to 1.7% BF for BOD POD − DXA, are likely due in part to differences in laboratory equipment, study design, and subject characteristics and in some cases to failure to follow the manufacturer's recommended protocol. Wide intersubject variations between methods are partly attributed to technical precision and biological error but to a large extent remain unexplained. On the basis of this review, future research goals are suggested. *Am J Clin Nutr* 2002;75:453–67.

KEY WORDS Body-composition methods, air-displacement plethysmography, hydrostatic weighing, dual-energy X-ray absorptiometry, isotopic dilution, total body water, multicompartment body-composition models, thoracic gas volume, residual lung volume, review

INTRODUCTION

Air-displacement plethysmography has been used to measure human body composition for nearly a century, but was not developed into a viable system for routine use until the mid-1990s (1). There is only one commercially available system for air-displacement plethysmography, which is known by the trade name BOD POD (Life Measurement, Inc, Concord, CA). Air-displacement plethysmography offers several advantages over established reference methods, including a quick, comfortable, automated, noninvasive, and safe measurement process, and accommodation of various subject types (eg, children, obese, elderly, and disabled persons). However, as with any new body-composition technology, it is important to establish its validity, reliability, and practicality in various populations.

In this review, we summarize the principal findings from studies published between December 1995 (the time at which the BOD POD was initially validated) and August 2001 that compared the BOD POD with reference methods. Specifically, we compared in both adults and children the reliability and validity of the BOD POD with the reliability and validity of established reference methods, ie, hydrostatic weighing (HW), dual-energy X-ray absorptiometry (DXA), and multicompartment [3-compartment (3C) and 4-compartment (4C)] models. To fully comprehend the significance of the viability of the BOD POD today, it is necessary to gain an understanding of the history of the development of air-displacement plethysmography. Therefore, we provided a brief description and historical overview of air-displacement plethysmography in general and of the BOD POD in particular and reviewed the operating principles of the BOD POD. Finally, we discuss the potential applicability of air-displacement plethysmography for use in a wide range of populations and summarize areas in need of further research.

[1] From the Department of Internal Medicine, the Center for Human Nutrition, Washington University, St Louis (DAF); the Department of Preventive Medicine, the Institute for Preventive Research, the Keck School of Medicine, the University of Southern California, Los Angeles (MIG); and the Energy Metabolism Laboratory, the Jean Mayer US Department of Agriculture Human Nutrition Research Center on Aging, Tufts University, Boston (MAM).

[2] The contents of this publication do not necessarily reflect the views or policies of the US Department of Agriculture.

[3] Address reprint requests to MA McCrory, Energy Metabolism Laboratory, Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, 711 Washington Street, Boston, MA 02111-1524. E-mail: mmccrory@hnrc.tufts.edu.

Received June 12, 2001.
Accepted for publication August 31, 2001.

BACKGROUND AND BRIEF HISTORICAL PERSPECTIVE

Plethysmography refers to the measurement of size, usually volume. In addition to air-displacement plethysmography (1), there are several other techniques for measuring whole-body volume. These techniques include acoustic plethysmography (2, 3), helium displacement (4, 5), photogrammetry (6), and more recently, 3-dimensional photonic scanning (7) and sulfur hexafluoride dilution (8). However, this review is limited to a discussion of air-displacement plethysmography.

In air-displacement plethysmography, the volume of an object is measured indirectly by measuring the volume of air it displaces inside an enclosed chamber (plethysmograph). Thus, human body volume is measured when a subject sits inside the chamber and displaces a volume of air equal to his or her body volume. Body volume is calculated indirectly by subtracting the volume of air remaining inside the chamber when the subject is inside from the volume of air in the chamber when it is empty. The air inside the chamber is measured by applying relevant physical gas laws. Boyle's Law states that at a constant temperature, volume (V) and pressure (P) are inversely related:

$$P_1/P_2 = V_2/V_1 \qquad (1)$$

Therefore, when a constant temperature is maintained (isothermal conditions), Boyle's Law can be applied. Consequently, most early plethysmographs required temperature-controlled surroundings and isothermal conditions within the test chamber. This presented burdensome requirements for testing conditions, which restricted practical implementation of air-displacement plethysmography. As discussed later, this problem was not fully resolved until systems were developed that do not require isothermal testing conditions (1, 9, 10).

The principles of plethysmography were first applied to the measurement of the body volume and composition of infants in the early 1900s (11, 12), but it was not until the 1960s that relatively stable measurements were achieved (13, 14). However, these systems required that ambient conditions be maintained constant. Therefore, to deal with rapid fluctuations in temperature, humidity, and pressure generated by humans inside the enclosed chamber, the measurement process by necessity included procedures that were difficult and laborious and by modern standards would be considered impractical and unacceptable. For example, the infant plethysmograph developed by Friis-Hansen (13) needed to undergo a 1–2-h calibration procedure before each measurement, and the test procedure took an additional 2–3 h. The technique also required that a plastic catheter be inserted through the infant's nose into the stomach to achieve a direct connection between the air inside the infant and the air in the surroundings. Another example in which extreme measures were necessary is the use of Gundlach and Visscher's adult plethysmograph (14). This procedure required that the test chamber be filled with polyurethane foam to maintain isothermal conditions. In addition, the adult subject had to be wrapped in a goose-down blanket and was required to hold his or her breath for ≈10 s during the measurement. Because of inconveniences such as these and various technologic difficulties, none of the early air-displacement plethysmographs were ever developed for common, everyday use.

Later experimental air-displacement plethysmographs developed in the 1980s were more advanced technologically. Petty et al (9) used a motor-driven pump and oscillating piston to create pressure changes within their system designed for adults; they also used advanced electronics and material to absorb moisture build-up in the chamber during the 5-min test period. An infant plethysmograph developed by Taylor et al (10) used a 2-chambered, dynamic, pressure-differential system. Pistons between the 2 chambers moved in concert and were controlled by a sinusoidal crank. A high-pass filter (controlled leak) was also incorporated, and harmonic analysis was done to interpret the pressure signal. Despite major improvements over previous systems, the results from these newer systems were still not sufficiently accurate and repeatable for routine human body-composition measurements.

BASIC PRINCIPLES OF THE BOD POD

In the mid-1990s, the BOD POD became the first commercially available air-displacement plethysmograph. The physical design and operating principles of this system are described in detail elsewhere (1, 15) and are summarized here. The BOD POD system includes the BOD POD plethysmograph, electronic weighing scale, calibration weights and cylinder, computer, and software. The BOD POD is functionally divided into 2 chambers: a test chamber (for the subject) and a reference chamber. The internal volumes of these chambers are ≈450 and 300 L, respectively. A diaphragm oscillates between the chambers, producing sinusoidal volume perturbations that are equal in magnitude but opposite in sign. The perturbations result in very small pressure changes within the chambers (±≈1 cm water), which are monitored by transducers and analyzed for pressure at the frequency of oscillation (3 Hz). The ratio of the pressures is a measure of the test chamber volume. Unlike with early air-displacement plethysmographs, it is not necessary to conduct measurements under isothermal conditions in the BOD POD. Instead, the air in the chambers is allowed to compress and expand adiabatically (ie, it freely gains and loses heat during compression and expansion). In this case, the BOD POD makes use of Poisson's Law, which describes the pressure-volume relation under adiabatic conditions:

$$P_1/P_2 = (V_2/V_1)^{\gamma} \qquad (2)$$

where γ is the ratio of the specific heat of the gas at constant pressure to that of constant volume and is equal to 1.4 for air (16).

Although body-volume measurements in the BOD POD occur under mostly adiabatic conditions, there is some volume of air maintained under isothermal conditions that must be taken into account. The reason for this is that when there are small changes in pressure, isothermal air volumes are compressed 40% more than are adiabatic air volumes. The largest sources of isothermal air are those contained in the lungs, near skin or hair, and in clothing. Isothermal air from clothing and hair on the head are minimized by having the subject wear a tight-fitting swimsuit and swim cap. (The manufacturer of the BOD POD recommends the use of swimsuits and caps made from either Lycra (DuPont, Wilmington, DE) or other spandex-type material, for reasons discussed later.) The average amount of air in the lungs during normal tidal breathing, thoracic gas volume (V_{TG}), is measured with the procedure described below. Alternatively, V_{TG} can be predicted. Finally, the effect of isothermal air near the skin's surface is estimated by calculating a surface area artifact (SAA). The SAA is automatically computed by the BOD POD's software as

$$\text{SAA (L)} = k\ (\text{L/cm}^2) \times \text{BSA (cm}^2) \qquad (3)$$

where k is a constant (derived empirically by the manufacturer; 1) and BSA is body surface area calculated from body weight and height with use of the formula by DuBois and DuBois (17). The SAA is typically ≈−1.0 L for average-sized adults; it is negative because it represents the apparent negative volume produced by the isothermal air space near the skin's surface. Note that if the wrong height is entered into the BOD POD software, the calculation of %BF will be in error because of inappropriate estimates of SAA. For example, for a 70-kg average-sized person, a 25-cm (≈10-in) error in height [168 cm (66 in) instead of 193 cm (76 in)] will result in a miscalculation of SAA (via BSA) and an error in body fatness of ≈0.5–0.7% BF. Thus, all of the BOD POD results should be routinely screened to determine whether any software data entry errors were made.

The measurement of body volume involves 3 steps. The first step is a standard 2-point calibration process: first with the chamber empty to establish baseline and then with a calibration cylinder (≈50 L) to establish range (duration: 50 s each). In the second step, the subject's volume in the chamber is measured (duration: 50 s). At this point, the measured body volume is "raw" (Vb_{raw}), ie, it has not been corrected for V_{TG} and SAA. This step is then repeated to check for agreement. If these 2 Vb_{raw} measurements are within 0.2% or 150 mL, whichever is larger, they are averaged. If the first 2 Vb_{raw} measurements do not meet these criteria, a third Vb_{raw} determination is made and the 2 values that are closest and within the criteria for agreement are averaged. If ambient conditions are relatively stable and the subject is breathing quietly in a relaxed fashion, it is common for the 2 Vb_{raw} measurements to agree within the predefined criteria. If the criteria are not met, the manufacturer suggests that the entire procedure be repeated, including the 2-point calibration step. Situations that could cause nonagreement between individual Vb_{raw} measurements include changing environmental conditions, other environmental impositions (eg, pressure changes in the room due to opening and closing doors or air drafts), or irregular tidal breathing by the subject (eg, yawning, throat clearing, or breath-holding). In the third step, V_{TG} is measured with the use of a procedure similar to that used in standard pulmonary plethysmography, sometimes called the panting maneuver by respiratory physiologists (18).

In contrast with traditional pulmonary plethysmography in which V_{TG} is determined at end-tidal exhalation [ie, functional residual capacity (FRC)], the BOD POD measures V_{TG} at midtidal exhalation. This is done because it is necessary to correct Vb_{raw} for the average amount of air in the lungs during normal tidal breathing, which is reflected by taking the measurement at midtidal exhalation. (A key assumption is that the subject is breathing normally during both the Vb_{raw} measurement and the V_{TG} measurement.) Thus, V_{TG} values derived from the BOD POD should be directly compared with V_{TG} values derived from a pulmonary plethysmograph only after correction for this difference (eg, a difference of ≈50% of the tidal volume). The V_{TG} measurement procedure begins with the subject breathing room air quietly through a disposable tube and antimicrobial filter while wearing a nose clip. After a few normal tidal breaths, a shutter valve in the airway closes, occluding it for ≈2 s. During occlusion, the subject makes 2 or 3 gentle quick puffs by alternately contracting and relaxing the diaphragm (ie, the panting maneuver). This leads to small changes in the gas volume of the airways, simultaneously with changes in body volume that are equal but opposite. These volume changes produce pressure changes that are monitored throughout the procedure. Comparison of the magnitudes of the changes in airway and chamber pressure allows calculation of V_{TG} via proprietary methods (Life Measurement, Inc, personal communication, 2001).

Two indicators are used to assess good compliance with the V_{TG} procedure: the figure of merit and airway pressure. The figure of merit is an index that estimates the degree of agreement between pressures measured inside the chamber and in the breathing airway (after scaling and translation). A smaller merit value indicates better agreement. Situations that may lead to poor agreement in these pressure values include lack of a tight lip seal around the tube, failure to wear a nose clip, significant puffing of the cheeks, or contraction of the abdominal muscles. Calculation of the figure of merit is discussed in detail by Dempster and Aitkens (1). If the airway pressure is too high, it may indicate closure of the glottis (ie, a Valsalva maneuver) or significant alveolar compression; both of these factors would result in falsely low V_{TG} values. If the figure of merit is >1.0 or the airway pressure is ≥35 cm water, the manufacturer recommends that the V_{TG} value be rejected and the procedure be repeated.

The BOD POD also allows for the prediction of V_{TG}. This feature is useful when it is necessary to test many subjects in a short period of time. Predicted V_{TG} was used in some studies when subjects were not able to satisfactorily perform the V_{TG} measurement procedure (19–21). The V_{TG} prediction equations currently used by the BOD POD (software version 1.69; Life Measurement, Inc) are based on FRC predictions by Crapo et al (22) from the heights and ages of subjects aged 17–91 y and include a further estimate for 50% of tidal volume. The accuracy of predicted V_{TG} and the effect of its use instead of measured V_{TG} on body-composition measurements are discussed below. Body volume in the BOD POD is calculated with the following formula:

$$Vb_{corr}\ (L) = Vb_{raw}\ (L) - SAA\ (L) + 40\%\ V_{TG}\ (L) \qquad (4)$$

where Vb_{corr} is the body volume corrected for SAA and V_{TG}. As part of the test procedure, the subject is also weighed to the nearest gram on the BOD POD's electronic scale. The provided calibration weights allow the operator to calibrate the scale periodically to ensure accuracy. Once body mass (M) and Vb_{corr} are known, the principles of densitometry are applied (23, 24). Body density (D_b) is calculated as M/Vb_{corr}, and D_b is then inserted into a standard formula for estimating %BF based on a 2-compartment (2C) model, such as the models of Siri (24) or Brozek et al (25) for whites and of Schutte et al (26) or Wagner and Heyward (27) for blacks. Alternatively, D_b can be used in multicompartment models (eg, 3C and 4C models) as discussed later.

RELIABILITY OF THE BOD POD

Reliability is a general term denoting repeatability or consistency between ≥2 measurements. The reliability of the BOD POD in different studies has been reflected by many statistical terms, such as SD, CV, precision (*see* definition below), intraclass correlation, and mean differences between tests. For the purposes of this review, we chose to limit the discussion of the BOD POD's reliability to only the most consistently reported statistics: SD, CV, and precision (defined as $[(SD/n)/\sqrt{d}]$, where n is the sample size and d is the number of repeated measurements).

Inanimate objects

The reliability of the BOD POD in measuring the body volume of inanimate objects is reported to be excellent. Twenty consecutive measurements of a 50.039-L aluminum cylinder resulted in a

TABLE 1
Reliability of percentage body fat measured with the BOD POD in adults[1]

Reference	n	CV	Number of trials or days
Within day		%	
McCrory et al, 1995 (28)	16	1.7 ± 1.1[2]	2 trials
Iwaoka et al, 1998 (8)	7	3.7 ± 4.3	2 trials
Sardinha et al, 1998 (29)	NR	3.3[3]	2 trials
Biaggi et al, 1999 (30)	NR	2.3 ± 1.9[3]	2 trials
Miyatake et al, 1999 (31)	5	2.5 ± 0.8	2 trials
Miyatake et al, 1999 (31)	5	4.5 ± 5.8[4]	3 trials (different operators)
Between day			
Nuñez et al, 1999 (20)	4	2.0 ± 0.1	4 d
Miyatake et al, 1999 (31)	10	2.3 ± 0.9	3 d
Levenhagen et al, 1999 (32)	NR	2.0 ± 2.1[3]	7 d

[1] NR, not reported.
[2] $\bar{x}$ ± SD.
[3] Reported as unpublished observations in the discussion sections of these articles.
[4] Reduces to 2.7 ± 2.0% if one abnormal test result is discarded.

mean (±SD) volume of 50.027 ± 0.00127 L and a corresponding CV of 0.025% (1). Results were similar when the experiment was repeated on another day. In another study, repeated measurements over 4 d of smaller volumes ranging from 4.643 to 50.0 L resulted in a mean CV of 0.67 ± 0.70% (20).

Adults

Reliability of percentage body fat

Seven studies reported the reliability of %BF measured by the BOD POD (8, 20, 28–32) as CVs; these values are shown in **Table 1**. Reported mean within-subject CVs for %BF ranged from 1.7% to 4.5% within a day and from 2.0% to 2.3% between days. These CVs are within the range of those measured previously by HW (8, 28, 33, 34) and DXA (35–37). In the 2 studies that examined the within-day repeatability of the BOD POD and HW in the same subjects, CVs did not differ significantly between methods: 1.7% compared with 2.3% in the study by McCrory et al (28) and 3.7% compared with 4.3% in the study by Iwaoka et al (8). Miyatake et al (31) reported similar mean CVs for tests conducted on the same day and on different days (over 3 d). They also reported a mean intertester CV of 4.5% (3 different operators). Examination of the individual data showed that this unexpectedly high CV was due to one abnormal test result in 1 of 5 subjects measured by 1 of 3 operators and may have been an anomaly. [Note that Wells and Fuller (38) suggest routinely conducting 2 tests per subject, enabling detection of infrequent rogue BOD POD results such as these.] Recalculation of the mean CV without the abnormal test result gave a mean intertester CV of 2.7%. Further studies in different populations and with larger numbers of subjects are needed to determine usual values for within-day, between-day, and intertester CVs.

Reliability of body volume

Two groups of investigators examined the reliability of body-volume measurement by the BOD POD relative to that with HW in adults. Dewit et al (39) and Wells et al (7) both reported that the precision (defined above) of Vb_{corr} was better with the BOD POD (0.07 and 0.11 L, respectively) than with HW (0.15 and 0.16 L, respectively). It is important to point out that in both of these studies, V_{TG} was predicted rather than measured. In contrast, lung volume at submersion was measured in conjunction with HW (JCK Wells, personal communication, 2001). This use of a constant, albeit predicted, V_{TG} value would tend to bias the precision of the BOD POD toward a more consistent body-volume measurement compared with when the precision of HW is calculated with a measured and presumably variable lung volume. Future studies are needed to quantify the precision of the BOD POD when measured V_{TG} values are used; this will provide a more direct comparison with the precision of HW.

Children

Reliability of percentage body fat

The CV for repeated %BF measurements by the BOD POD in children has not been reported. Using the precision statistic described above, Wells and Fuller (38) described the precision of 2 repeat measurements of %BF to be 0.83% for 11 boys ($\bar{x}$: 12.6%) and 0.99% for 16 girls ($\bar{x}$: 19.7%). Precision was not related to body size because duplicate measurements in 30 men and women with 18.0% and 27.5% BF, respectively, had similar values for precision (0.99% and 0.76% BF).

Reliability of body volume

Dewit et al (39) and Wells et al (7) reported the precision of body-volume measurements in children aged 7–14 y. Precision of Vb_{corr} was 0.07 and 0.08 L in the 2 studies, respectively, which was just as good as or slightly better than the precision in adults in the same studies (0.07 and 0.11 L, respectively). Therefore, the precision of body-volume measurements in children and adults was comparable in these 2 studies, despite the smaller body volumes of the children. Similar body-volume precision was reported in another study by the same research group (38). It has been suggested that a relatively small ratio of chamber volume to subject volume would optimize the precision of body-volume measurements (5, 9). For example, Gnaedinger et al (5) calculated a mean ratio of chamber volume to subject volume of 6:1 in their plethysmograph and suggested that a smaller ratio would have improved their data. Assuming a BOD POD test chamber volume of 450 L, the mean ratio of chamber volume to subject volume can be calculated from data provided by Dewit et al (39). Despite the larger ratio for children (14:1 for children compared with 8:1 for adults), the precision of measurements in children and adults was similar. This finding indicates that within the range of body sizes studied thus far, the ratio of chamber volume to subject volume may be irrelevant in the BOD POD.

VALIDITY OF THE BOD POD RELATIVE TO HW

Summary of findings in adults

A summary of studies that compared body-composition measurements by the BOD POD and HW in adults is shown in **Table 2**. Most of these studies were conducted in young to middle-aged subjects (age range: 20–56 y), except for the study by Nuñez et al (20), which included subjects ≤86 y of age. BMI ranged from 17 to 40 across the different studies.

Mean group differences between the BOD POD and HW measurements ranged from −4.0% to 1.9% BF; 5 of the 12 studies showed no significant differences between the 2 methods (7, 8, 19, 20, 28, 30, 32, 39–43). Of the 7 studies that did show

TABLE 2
Summary of studies that compared percentage body fat (%BF) measurements made with the BOD POD or hydrostatic weighing (HW)[1]

Reference	Number of subjects	Sex	Age[2]	BMI[2]	BOD POD − HW[3]	Regression analysis[4] Slope	R^2	SEE	Bland-Altman analysis 95% Limits of agreement	Significant trend	Significant sex effect
	n		*y*	*kg/m²*	*%BF*						
Adults											
McCrory et al, 1995 (28)	68	M,F	20–56	18–36	−0.3 ± 1.6	0.94	0.93	1.8	−4.0, 3.4	No	No
Iwaoka et al, 1998 (8)	7	M	31–44	22 ± 4	−4.0 ± 3.1[5]	0.78[6]	0.82	NR	−10.1, 2.2	No	NA
Biaggi et al, 1999 (30)	47	M,F	19–48	NR	−0.1	0.82	0.89	NR	−6.1, 5.9	Upward	Yes
Collins et al, 1999 (19)[7]	69	M	19 ± 1	NR	−2.0[5]	0.91[8]	0.89	2.2	−6.4, 2.5	No	NA
Levenhagen et al, 1999 (32)	20	M,F	19–47	20–36	−0.5	0.77[8]	0.94	NR	−6.7, 5.7	Upward	Yes
Nuñez et al, 1999 (20)[7]	72	M,F	20–86	≈25 ± 4	0.1	NR	0.90[9]	NR	NR	NR	No
Dewit et al, 2000 (39)[7]	10	M,F	19–41	21 ± 2	−3.3 ± 2.3[5]	NR	NR	NR	NR	NR	NR
Fields et al, 2000 (40)	67	F	18–55	17–34	1.2 ± 2.3[5]	0.96	0.94	2.3	−3.5, 5.8	No	NA
Wagner et al, 2000 (41)	30	M	19–45	19–40	1.9[5]	NR	0.84[9]	NR	NR	No	NA
Wells et al, 2000 (7)[7]	22	M,F	31 ± 8	22 ± 3	−2.2 ± 3.3[5,10]	NR	NR	NR	NR	NR	NR
Millard-Stafford et al, 2001 (42)	50	M,F	25 ± 6	≈24	−2.8 ± 4.1[5]	0.76[6]	0.78[9]	NR	−11.0, 5.4	Upward	NR
Fields et al, 2001 (43)	43	F	19–54	17–37	0.2 ± 2.4	0.90	0.94	2.3	−4.9, 5.1	No	NA
Children											
Nuñez et al, 1999 (20)[7]	48	M,F	6–19	≈21 ± 4	1.2	NR[8]	0.83[9]	NR	NR	NR	NR
Dewit et al, 2000 (39)[7]	22	M,F	8–13	17 ± 2	0.8 ± 5.4	NR	NR	NR	NR	NR	NR
Fields and Goran, 2000 (44)[11]	25	M,F	9–14	13–35	2.6 ± 3.4[5]	0.86	0.87	3.3	−4.4, 9.6	No	No
Lockner et al, 2000 (21)[7]	54	M,F	10–18	NR	−2.9[5]	NR	0.72	NR	NR	NR	NR
Wells et al, 2000 (7)[7]	10	NR	7–14	17 ± 2	0.6 ± 0.7[10]	NR	NR	NR	NR	NR	NR

[1] All studies used Siri's equation (24) to convert body density to %BF, with the following exceptions: reference 10 used Brozek et al's equation (25), references 30 and 32 used Schutte et al's equation (26) in blacks, reference 41 used Schutte et al's equation (26) and Wagner and Heyward's equation (27) in blacks, references 7 and 39 used the child-specific equation developed by Wells et al (45), and reference 44 used child-specific equations developed by Lohman (46). NR, not reported; NA, not applicable.
[2] Range, or $\bar{x} \pm$ SD when range was not reported.
[3] Difference ($\bar{x}$ or $\bar{x} \pm$ SD) in %BF between the 2 methods.
[4] Prediction of %BF with HW from %BF measured with the BOD POD.
[5] Significantly different from 0, $P < 0.05$.
[6] Statistical comparison of the slope with 1.0 was not reported.
[7] Some or all BOD POD tests were done by using predicted thoracic gas volume; Wells et al (7) used child-specific equations for predicting thoracic gas volume in children.
[8] Significantly different from 1.0, $P < 0.05$.
[9] From regression analysis using body density rather than %BF.
[10] Statistical significance not reported in the original articles; however, the P value = 0.008 for adults and 0.11 for children (JCK Wells, personal communication, 2001).
[11] Data for fat mass were originally reported, but were recomputed for this review in %BF units with the use of Siri's equation (24) to facilitate comparison with other studies.

a significant mean difference, the direction of the differences was inconsistent: 5 (7, 8, 19, 39, 42) showed a lower %BF with the BOD POD than with HW and 2 (40, 41) showed the opposite. Note that the largest mean differences (−4.0% and −3.3% BF) occurred in the 2 studies that had the fewest subjects ($n \leq 10$) (8, 39). Ethnicity did not contribute significantly to differences between the 2 methods in the 2 studies that had a wide enough range of ethnicities to examine this possibility (20, 28); however, the potential effects of ethnicity were not reported in 2 studies that included both whites and blacks (19, 42).

In the 8 studies that reported regression analysis for the prediction of %BF measured by HW from %BF measured by the BOD POD, the slope of this relation ranged from 0.76 to 0.96; the mean value was much lower than the desired value (1.00) in 4 of these studies (8, 30, 32, 42). Not all of the studies reported whether this slope differed significantly from 1.00; of those that did (19, 28, 30, 32, 40, 43), only 2 studies (19, 30) had slopes that differed significantly from 1.00, as indicated in Table 1. %BF measured by the BOD POD explained 78–94% of the variance in %BF measured by HW, whereas the SEEs reported in 4 of the 12 studies ranged from 1.8% to 2.3% BF. These SEEs are in the excellent to ideal range (≤2.5 %BF) according to Lohman (47).

Bland-Altman limits of agreement (mean difference ± 2 SD ranges; 48) and results of trend analysis are also shown in Table 2. In general, the limits of agreement indicated wide variations in agreement between the BOD POD and HW (range: ≈9–16% BF) for individuals, even when group mean differences were small.

Summary of findings in children

Relatively few studies have compared the BOD POD with HW in children (Table 2). Of the 5 studies that have (7, 20, 21, 39, 44), the age range of the children studied was 6–19 y. Two of these studies (21, 44) reported that, on average, the BOD POD gave significantly different %BF measurements than did HW. As in the studies in adults, the results were in opposite directions (2.6 compared with −2.9% BF, respectively). The other 3 studies (7, 20, 39) reported that %BF measured by the BOD POD was somewhat higher than that measured by HW (0.6–1.2% BF), but not significantly so. The slope of the relation for the prediction of %BF by HW from %BF by the BOD POD was 0.86, which was not significantly different from 1.00 in the one study that reported the slope (44). In the 3 studies that reported R^2 values, the BOD POD explained between 72% and 87% of the variation in HW (20, 21, 44). The only SEE available (3.3% BF) was from Fields and Goran (46), which was in the good (average) range (47). Finally, Bland-Altman limits of agreement calculated from the study by Fields and Goran (44) were −4.4% to 9.6% BF, indicating large individual variations in the difference between the BOD POD and HW.

Potential reasons for differences between the BOD POD and HW measurements

Theoretically, the BOD POD and HW should give identical values for D_b and %BF because both methods are based on the principles of densitometry. Therefore, any differences between the 2 methods can be attributed to differences in either measured body mass (if the same scale is not used for both methods) or body volume. In turn, differences in body volume measured with the BOD POD can be attributed to variations in measurements of Vb_{raw}, SAA, or V_{TG}, and differences in body volume measured with HW can be attributed to variations in body mass measured in water, residual lung volume (V_R), or other types of lung volume [eg, lung volume at submersion (7, 39)].

Interlaboratory variation

Interlaboratory variation may be an important factor contributing to the discrepant findings among studies in mean differences between the BOD POD and HW. The extent to which different BOD POD systems vary is not known, although it is hypothesized that BOD POD systems may vary less than do HW systems because there are several variations of HW equipment and methods (eg, different weighing scales and methods for measurement of lung volume) but only one type of BOD POD system manufactured by one company. Although it is possible that the variation in mean differences in the previously mentioned adult studies was random, note that there are 4 pairs of studies, with each of the 4 pairs being from a different laboratory but with each study within a pair being from the same laboratory [(7) and (39), (30) and (32), (40) and (43), and (19) and (42)], and the results within each of the study pairs are more similar than among the study pairs. For example, Dewit et al (39) and Wells et al (7) reported large negative mean differences between the 2 methods (−3.3% and −2.2% BF, respectively), as did Collins et al (19) and Millard-Stafford et al (42) (−2.0% and −2.8% BF, respectively). However, Biaggi et al (30) and Levenhagen et al (32) reported 2 of the smallest and slightly negative mean differences (−0.1% and −0.5% BF, respectively) and Fields et al (40, 43) reported mean differences that were slightly positive, with one value being close to 0 (1.2% and 0.2% BF, respectively). These similar findings within study pairs suggest that interlaboratory variation in protocol, test equipment, or both may contribute importantly to the variation in results observed among studies. To more fully understand the potential effect of interlaboratory variation on measurements of %BF, a multicenter study in which the same subjects are tested in different laboratories is needed.

Test conditions

Measurements with the BOD POD should be made under standard test conditions, ie, subjects should wear minimal but skintight clothing [Lycra (DuPont) or other spandex-style swimsuit and cap], be completely dry, and be in a resting state.

Effects of clothing. In some of the studies discussed, subjects wore spandex-style shorts (rather than swimsuits, which are recommended by the manufacturer) while undergoing measurements with the BOD POD. This may have contributed to the relatively lower %BF values measured with the BOD POD than with HW in some of the studies (19, 21). In other studies it is unclear what type of clothing was worn during the test protocol. However, it is known that excess clothing causes a significant underestimation of body volume because air that comes in contact with cloth will remain isothermal as pressure fluctuates. The more cloth that is worn, the larger the layer of isothermal air. Because isothermal air is 40% more compressible than is adiabatic air, body volume (Vb_{raw}, and hence Vb_{corr}) is underestimated and, in turn, D_b is overestimated and %BF is underestimated. The effect of excess clothing on %BF measurements with the BOD POD was illustrated in a study by Fields et al (40). No significant difference in %BF was found between women who wore a 1-piece or 2-piece swimsuit. However, %BF was ≈5% lower in women who wore a hospital gown than in women who wore either type of swimsuit. Although this study illustrated that extreme deviations from the manufacturer's recommended protocol (ie, wearing of loose clothing) had significant effects on estimates of %BF with the BOD POD, it did not address whether slight deviations from the recommended protocol (ie, wearing of spandex-style shorts rather than a swimsuit) would result in acceptable %BF measurements. Until studies are conducted that confirm or deny that alternative clothing is acceptable, it is suggested that the clothing protocol recommended by the manufacturer be rigorously followed.

Effects of testing under nondry, nonresting conditions. In 2 studies (21, 32), the order in which the 2 methods were conducted was randomized; therefore, in some cases the BOD POD measurements were made first and in others the HW measurements were made first. However, neither of these studies reported whether the subjects were still wet when the BOD POD measurements were made or how much time passed between the 2 tests. Tests with the BOD POD should be conducted only when the subjects are completely dry and in a rested state. Moisture on the body, in the hair, and in the swimsuit will artificially increase body weight. Furthermore, if subjects are recovering from situations that elevate metabolism (eg, exercise or presence in a tank of warm water for 10–15 min as part of the HW procedure), breathing patterns are likely to change over time. In BOD POD testing, a key assumption is that breathing patterns are similar during the Vb_{raw} and V_{TG} measurements; however, this will not be the case if subjects are recovering from a physical stress. This situation is somewhat analogous to HW when V_R is measured on land and it is assumed that the subject exhales to the same end point both on land and in the water. In both cases, the exact lung volume is not a concern, but the lung volume should be the same during the HW and V_R measurement procedures and, likewise, during the Vb_{raw} and V_{TG} measurement procedures.

The effect of testing under nondry, nonresting conditions was illustrated in a preliminary study (DA Fields, GR Hunter, unpublished observations, 2000). When the BOD POD tests were conducted 10–15 min after HW, BF was 2.3% lower than it was when measured before HW. In that study, subjects had dried with a towel after HW but their hair and swimsuits were still damp when the BOD POD measurements were made.

V_{TG} prediction

In some studies, predicted V_{TG} was used when some subjects could not adequately perform the panting maneuver to obtain

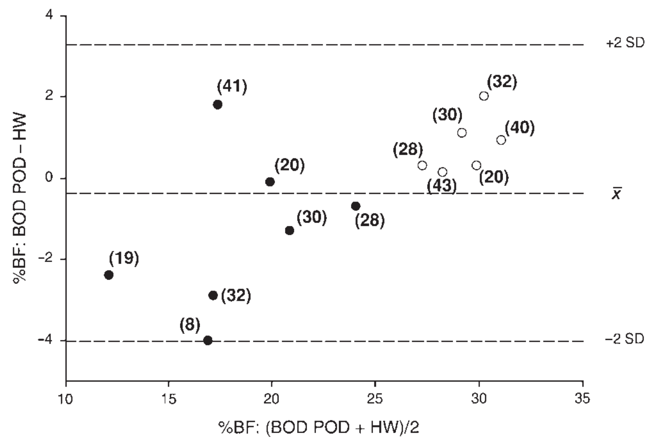


FIGURE 1. Bland-Altman plot of sex-specific mean differences between percentage body fat (%BF) measured with the BOD POD (Life Measurement, Inc, Concord, CA) and with hydrostatic weighing (HW) in men (●) and women (○) in individual studies (reference numbers in parentheses). For reference 19, the subsample that was also tested by DXA was used. Only studies published from December 1995 to August 2001 that provided mean data for men and women separately were used. The relation between the difference between the 2 methods and the average of the 2 methods was significant ($r = 0.66$, $P = 0.014$).

measured V_{TG} (20, 21), whereas in others (7, 39) it was used routinely simply to save time (JKC Wells, personal communication, 2001). McCrory et al (49) reported no significant difference between mean predicted and measured V_{TG} in 50 men and women aged 18–56 y (BMI: 19–35) with the use of software versions 1.50 and 1.53 (Life Measurement, Inc). Further findings indicated that for 82% of the subjects, the use of predicted V_{TG} resulted in a value within ±2% BF of that calculated with the use of measured V_{TG}. The difference between predicted and measured V_{TG} was not related to the magnitude of V_{TG} (MA McCrory, PA Molé, TD Gomez, KG Dewey, EM Bernauer, unpublished observations, 1998). In contrast with the results of the above study, 2 later studies (software version not reported) showed that, on average, predicted V_{TG} was significantly higher than measured V_{TG}, by 344 mL in 69 collegiate football players (19) and by 190 mL in 37 children aged 10–18 y (21). These findings suggest that the BOD POD's current method for prediction of V_{TG} may not be valid for all populations and illustrate the need to report software versions used in all studies to help determine whether different versions of software may be responsible for conflicting findings among studies.

Errors in V_{TG} prediction generally have only a small effect on %BF. As can be deduced from Equation *4*, overestimation of V_{TG} results in overestimation of Vb_{corr} and, in turn, underestimation of D_b and overestimation of %BF. However, because only 40% of V_{TG} is incorporated into the equation to calculate Vb_{corr}, the magnitude of the overestimation of V_{TG} reported in the above studies should only have caused a very small overestimation of %BF (<1.0%). Note that, in the studies by Collins et al (19) and Lockner et al (21), %BF measured with the BOD POD was significantly lower (rather than higher as would be caused by overprediction of V_{TG}) than that by HW (Table 2). This finding indicates that other factors (eg, clothing) may have contributed to the observed differences between the BOD POD and HW measurements.

The study by Lockner et al (21) indicates that some children may have more difficulty performing the V_{TG} procedure than do adults; only 69% of their study population adequately performed the V_{TG} measurement procedure in 3 trials. In contrast, Fields and Goran (44), who studied children of a similar age range, obtained V_{TG} measurements in all of their subjects. Valid measurements, on the basis of the standard merit and airway criteria, were obtained in ≈80% of the children in ≤ 3 trials and in ≈20% of the children in >3 trials. Two studies conducted in children aged 5–14 y (39, 45) substituted child-specific prediction equations for FRC (50) and tidal volume (51) to calculate child-specific V_{TG} and body-composition measurements with the BOD POD. In these studies, neither measured nor predicted V_{TG} with the BOD POD was reported; therefore, it is not possible to assess the utility of these child-specific prediction equations. However, Dewit et al (39) noted that when the child-specific equations were used, rather than the adult equations that were incorporated into the BOD POD's software, the mean difference in %BF (calculated as BOD POD − HW) changed from 0.8% to −0.9% BF. This finding suggests that the use of the adult equations overpredicts V_{TG} in children. This is understandable because the BOD POD was originally designed for use in adults. Although usual errors in V_{TG} have only a relatively small influence on %BF as discussed above, more work is needed to improve both the V_{TG} measurement process and the accuracy of V_{TG} prediction in different populations.

Subject sex

Whether the sex of the subject systematically affects the results obtained with the BOD POD or HW remains to be determined. This possibility was first raised by Biaggi et al (30), who reported a significant sex effect and found that the mean difference between the BOD POD and HW was positive for females (1.0 ± 2.5% BF) and negative for males (−1.2 ± 3.1% BF). The same research group also reported findings similar to those of Levenhagen et al (32). However, an additional 2 studies that included both males and females and that examined whether there was a significant effect of sex on the difference between %BF measured by HW and with the BOD POD found no effect of sex (20, 28). Additionally, the studies by Biaggi et al (30) and Levenhagen et al (32) were the 2 of the only 3 studies to report a significant upward trend in the Bland-Altman plot (Table 2), indicating a negative difference between the BOD POD and HW measurements in leaner subjects and a positive difference in fatter subjects. Millard-Stafford et al (42) also reported a significant upward trend, but did not specifically test for a sex effect, possibly because of the relatively small number of females in their study (10 females and 40 males). Because males tend to be leaner than females, it is difficult to determine whether the significant effect of sex reported in the studies by Biaggi et al (30) and Levenhagen et al (32) were due to an effect of sex per se or to body fatness. Examination of the Bland-Altman plots from these studies showed little overlap in %BF between men and women, although it would be possible in future studies to recruit men and women matched for %BF in an attempt to disentangle the separate influences of %BF and sex.

To further examine the question of whether differences between the BOD POD and HW measurements are dependent on the sex of the subject or on %BF, we plotted the sex-specific means in Bland-Altman fashion for studies in which mean differences were reported separately for males and females (8, 19, 20, 28, 30, 32, 40, 41, 43) (**Figure 1**). An upward trend was seen ($r = 0.66$, $P = 0.014$), with no overlap in mean %BF between males and females. Therefore, in this analysis, as in the individual studies, it is impossible to separate the confounding effects of subject sex and %BF.

Biaggi et al (30) hypothesized that the sex effect observed in their study may have been attributable to the greater amount of body hair on men than on women. Theoretically, excess body hair may reduce apparent body volume by increasing the amount of isothermal air near the surface of the body as explained above. Thus, body volume may be underestimated if more isothermal air than usual is present next to the skin, remaining unaccounted for by the BOD POD's SAA estimation. In fact, the effect of animal fur on air-displacement plethysmography measurements was shown in 1985 by Taylor et al (10), who found that the measured volume of rats was 15% lower by air-displacement plethysmography than by HW; volume was not underestimated in inanimate objects.

It is possible that body hair on humans does not routinely influence the accuracy of body-volume measurements, except in subjects who have an unusually thick layer of body hair and that the men in Biaggi et al's study (30) were unusually hairy. To definitively answer the question of whether body hair significantly influences air-displacement plethysmography measurements of body volume, a study is needed in which these measurements are conducted before and after the body is shaved. This was done to a limited extent in men (52). The men in the study grew beards for 3 wk and then the BOD POD measurements were made before and after the beards were shaved. Although there were large individual variations, mean Vb_{raw} was 157 mL lower and %BF was 0.9% lower after shaving. These findings suggest that for men who have beards, an additional factor could be built into the BOD POD software to adjust for the small effect of additional isothermal air associated with a beard. The findings further suggest that during longitudinal studies in which the BOD POD is used to measure body composition, men should either remain clean shaven or maintain the same amount of facial hair throughout the study.

Subject size

Lockner et al (21) reported that the difference between D_b by HW and the BOD POD in children was significantly related to height, body mass, and body surface area, with the largest differences (calculated as BOD POD – HW) seen in the smallest children. The suggestion that a smaller ratio of chamber volume to subject volume would improve measurement precision (5, 9) and the above-mentioned findings of Lockner et al suggest that body-volume measurements with the BOD POD may be less accurate in smaller children than in larger children. However, as discussed above, there were other possible confounding factors in Lockner et al's study. Furthermore, the possibility that smaller (younger) children may have had more difficulty complying with the requirements of the HW procedure should not be overlooked.

Fasting compared with postprandial conditions

It is known that gas in the stomach or intestine that is not accounted for leads to an underestimate of D_b and an overestimate of %BF when measured by HW. This can be seen in the following formula used to calculate D_b by HW:

$$D_b = M_{land}/[(M_{land} - M_{water})/D_w] - V_R - V_{GI} \tag{5}$$

where D_b is in kg/L, M_{land} is body mass on land in kg, M_{water} is body mass in water in kg, D_w is the temperature-specific water density in kg/L, V_R is in L, and V_{GI} is gastrointestinal gas volume in L. Investigators often use an average estimate of 0.100 L (53) for intestinal gas. This estimate may be appropriate under fasting conditions; however, under postprandial conditions (even as long as 3 or 4 h after a meal; 19, 40, 43), the amount of intestinal gas varies depending on the specific foods ingested (54, 55). Theoretically, air-displacement plethysmography will at least partially account for gas in the intestine during the measurement of Vb_{raw} (53, 56, 57), perhaps even as part of the measured V_{TG} if the gas is located above the diaphragm (eg, in the esophagus). Preliminary results by McCrory et al (58) showed that immediately after ingestion of a carbonated soft drink (355 mL, or 12 oz), %BF increased by 2.6% when measured by HW but increased by only 0.9% when measured with the BOD POD. A small increase in V_{TG} was also noted after ingestion of the carbonated soft drink, but there was no change in V_R.

Errors in V_R compared with errors in V_{TG}

The largest contributor to HW variability is the error in measuring V_R (59, 60). Depending on the measurement technique used, V_R can vary by as much as 300 mL and consequently affect %BF estimates by HW up to ≈4% (61). When V_R is measured on land, errors in HW also arise when there is a mismatch between the amount of air exhaled on land relative to that in water. Friedl et al (62) conducted HW measurements on 3 d in a single week. In one-half of the subjects, a learning effect on the maximal exhalation procedure under water was noted such that over time subjects exhaled a greater amount of air (and thus had a higher mass in water). In contrast, a concomitant change in the V_R measurements on land was not observed. This learning effect under water, but not on land, resulted in an average BF measurement that was 1% lower on the third day than on the first day in these subjects. It is possible that simultaneous determination of V_R and body mass in water would have alleviated the mismatch observed in Friedl et al's study (62); however, this solution is controversial because some studies suggest that the measurement of V_R in water by gas dilution may be underestimated because of pulmonary gas trapping (63, 64). As can be seen by comparing Equations *4* and *5* and as discussed by McCrory et al (49), an error in V_{TG} has less of an effect on measurements made with the BOD POD than an error in V_R of the same magnitude has on HW. However, the variability in V_{TG} relative to that in V_R has not yet been reported. In addition, the validity of V_{TG} measured with the BOD POD needs to be established. One way to do this is to compare V_{TG} measurements made with the BOD POD with those made by standard pulmonary plethysmography (considered by pulmonary physiologists as the gold standard method for measuring lung volume; 65–67), after correction for differences in tidal volume as discussed above.

VALIDITY OF THE BOD POD RELATIVE TO DXA

Summary of findings in adults

Nine studies compared body-composition measurements by DXA and the BOD POD in adults with BMIs ranging from 17 to 40 (19, 20, 29, 31, 32, 41–43, 68; **Table 3**). Most of these studies were conducted in young to middle-aged subjects, but 2 of the studies also included adults aged >55 y). Mean differences between %BF measured by the BOD POD and DXA varied widely. The differences in %BF were significant in about one-half of the studies conducted: negative (range: −2.0% to −3.0%) in 4 of the studies (19, 29, 32, 42) and positive (1.7 %BF) in 1 of the studies (41).

TABLE 3
Summary of studies that compared percentage body fat (%BF) measurements made with the BOD POD or dual-energy X-ray absorptiometry (DXA)[1]

Reference	Number of subjects	Sex	Age[2]	BMI[2]	BOD POD − DXA[3]	Regression analysis[4] Slope	R^2	SEE	Bland-Altman analysis 95% Limits of agreement	Significant trend	Significant sex effect
	n		*y*	*kg/m²*	*%BF*						
Adults											
Sardinha et al, 1998 (29)	62	M	31–46	19–35	−2.6 ± 2.6[5]	NR	0.86	NR	−2.6, 7.8	No	NA
Levenhagen et al, 1999 (32)	20	M,F	19–47	20–36	−3.0 ± 3.7[5]	0.99	0.88	NR	−4.4, 10.4	No	No
Collins et al, 1999 (19)[6]	20	M	20 ± 1	NR	−2.0[5]	1.02[7]	0.80	2.4	NR	NR	NA
Miyatake et al 1999 (31)	16	M,F	28 ± 7	21 ± 3	NR	NR	0.83	NR	NR	NR	NR
Nuñez et al, 1999 (20)[6]	72	M,F	20–86	≈25 ± 4	−0.4	0.91[7]	0.88	3.5	NR	No	No
Koda et al, 2000 (68)	721	M,F	40–79	23 ± 3	≈−0.1 ± 3.8	NR	0.78–0.81	NR	NR	NR	Yes
Wagner et al, 2000 (41)[8]	30	M	19–45	19–40	1.7[5]	NR	0.86	2.8	NR	NR	NA
Fields et al, 2001 (43)	43	F	19–54	17–37	0.6 ± 3.4	1.10	0.91	3.4	−6.1, 7.2	No	NA
Millard-Stafford et al, 2001 (42)	50	M,F	25 ± 6	≈24	−2.5 ± 3.7[5]	NR	NR	3.7	NR	NR	NR
Children											
Nuñez et al, 1999 (20)[6]	48	M,F	6–19	≈21 ± 4	−0.1	0.86[7]	0.81	4.0	NR	Upward[9]	No
Fields and Goran, 2000 (44)[10]	25	M,F	9–14	13–35	−3.9 ± 4.0[5]	1.02	0.85	4.1	−11.9, 4.1	No[11]	No[11]
Lockner et al, 2000 (21)[6]	54	M,F	10–18	NR	−2.1[5]	NR	0.88	3.4	NR	NR	NR

[1] All studies used a DPX-L (Lunar, Madison, WI) to measure DXA except studies (29, 31, 68) in which a Hologic QDR 1500 or 4500 (Waltham, MA) was used. All studies used Siri's equation (24) to convert body density to %BF, with the following exceptions: references 31 and 68 used Brozek et al's equation (25), reference 32 used Schutte et al's equation (26) in blacks, reference 41 used Schutte et al's equation and Wagner and Heyward's equation (27) in blacks, and reference 44 used child-specific equations developed by Lohman (46). NR, not reported; NA, not applicable.

[2] Range, or $\bar{x}$ ± SD when range was not reported.

[3] Difference ($\bar{x}$ or $\bar{x}$ ± SD) between %BF between the 2 methods.

[4] Prediction of %BF with DXA from %BF measured with the BOD POD.

[5] Significantly different from 0, $P < 0.05$.

[6] Some or all of the BOD POD tests were done with the use of predicted thoracic gas volume.

[7] Statistical comparison of the slope with 1.0 was not reported.

[8] Data were derived with the use of the equation of Wagner and Heyward (27) for the conversion of body density to %BF for blacks.

[9] Nonsignificant trend.

[10] Data for fat mass were originally reported, but were recomputed for this review in %BF units with the use of Siri's equation (24) to facilitate comparison with other studies.

[11] DA Fields, MI Goran, unpublished observations, 2000.

One additional study with a substantial sample size of 721 and an overall mean difference in %BF of −0.1% reported a significant negative mean difference (−1.3%) for females and a significant positive mean difference (1.2%) for males (68). In 3 of the 4 studies reporting regression analyses, prediction of %BF by DXA from %BF by the BOD POD resulted in slopes very close to 1.00, ranging between 0.99 and 1.02 (19, 32, 43); in the remaining study, the slope was somewhat lower, 0.91 (20). The amount of shared variance between the 2 methods ranged from 78% to 91%, whereas SEEs ranged from 2.4% to 3.5% BF [which were distributed among the good, very good, and excellent categories, as subjectively assessed by Lohman (47)]. The 95% limits of agreement ranged from 10% to 15% in the 3 studies that reported Bland-Altman analyses (29, 32, 43), indicating very large differences between these 2 methods in some individuals.

Summary of findings in children

The 3 studies conducted in children that compared %BF measurements made with the BOD POD and with DXA are also summarized in Table 3 (20, 21, 44). The children in these studies ranged in age from 6 to 19 y and all 3 studies included both boys and girls. In 2 of these studies (21, 44), a significant negative mean difference between the 2 methods was reported (−3.9% and −2.1% BF), but in the other study (20) there was almost no difference (−0.1% BF). The prediction of %BF with DXA from %BF with the BOD POD produced a slope of 1.02 in one study (44), but a lesser slope of 0.86 in another study (20). %BF measured with the BOD POD accounted for 81–88% of the variance in %BF measured by DXA as indicated by the R^2 value. The SEEs ranged from 3.4% to 4.1% BF, which are noted as fairly good or good by Lohman (47). A wide range of individual differences between the BOD POD and DXA measurements was indicated by Bland-Altman analysis, with 95% limits of agreement of −11.9% and 4.1% BF (44). In addition, Nuñez et al (20) reported a nonsignificant upward trend in their Bland-Altman plot, but Fields and Goran (44) found no such trend.

Potential reasons for differences between the BOD POD and DXA measurements

Many of the issues discussed above that may have contributed to the differences between the BOD POD and HW measurements also pertain to the observed differences between the BOD POD and DXA measurements, particularly the clothing worn during the BOD POD test, the order in which the different body-composition tests were conducted, and the prediction of V_{TG}. Other factors that also may be at play include limitations in DXA and errors in the assumptions inherent to the 2C models of densitometry, which are used in the BOD POD to calculate %BF. These additional factors and the potential sex effect on differences between the 2 methods are discussed below.

Limitations of the densitometric 2C model

The 2C model for converting D_b to %BF divides the body into components of fat mass and fat-free mass. Among the assumptions inherent to this model are that the densities of these 2 components are 0.9 and 1.1 kg/L, respectively, and that these densities do not vary among individuals or populations (47).

Although these numbers appear to be relatively accurate for the general population, it is known that the density of the fat-free mass can differ substantially from 1.1 kg/L for particular groups of individuals, such as the elderly, children, and blacks (47). In individuals and groups in whom density deviates from these assumptions, body-composition estimates based on the 2C model are in error; 2C models can be improved, however. For example, for children and adolescents, who have not yet matured chemically and whose fat-free mass has a greater proportion of water and lesser proportion of mineral, Lohman (46) used average estimates for the water and mineral proportions of the fat-free mass to derive age- and sex-specific 2C equations. These equations, applicable for persons aged 1–18 y, should improve group estimates of %BF by densitometry (including both the BOD POD and HW) and are preferable to the equations of Siri (24) or Brozek et al (25) for this age group. Data from Roemmich et al (69) support the use of these equations. They showed that in young adolescents, Lohman's equations resulted in a mean %BF estimate that was much closer to that derived by the gold standard 4C model (discussed below) than was %BF calculated with Siri's equation; however, individual errors were still high, as shown by the Bland-Altman limits of agreement.

The density of fat-free mass is influenced in large part by bone mineral because the density of bone is markedly higher than that of other components of the fat-free mass. Koda et al (68), who studied men and women aged 40–79 y, reported that the difference between the BOD POD and DXA %BF measurements observed in their study was inversely associated with the bone mineral content expressed as a percentage of the fat-free mass (BMC/%FFM) in both sexes. In other words, the lower the BMC/%FFM, the more positive the difference between the BOD POD and DXA %BF measurements. Of note, there are previous reports that BMC/%FFM is also inversely associated with differences between HW and DXA measurements (70–72). The mean BMC/%FFM in the study by Koda et al (68) was relatively low in both sexes [4.4% compared with Brozek et al's (25) estimate of 5.6% in a reference man]; on the basis of this information, it can be predicted from the 2C model that the BOD POD would overestimate %BF in both males and females in this age group. However, mean %BF measured with the BOD POD was significantly higher than that measured with DXA only in males. Furthermore, a large proportion of both sexes had a lower %BF as measured with the BOD POD (≈25% of males and 68% of females). This suggests that other factors in addition to a relatively low BMC/%FFM were responsible for the differences in %BF measured with the BOD POD and DXA. Koda et al (68) also reported that for both sexes, the differences between the BOD POD and DXA measurements of %BF (calculated as BOD POD – DXA) were positively associated with age, waist circumference, and sagittal diameter; ie, older subjects and those with larger waist circumferences and sagittal diameters had a more positive difference. However, multiple regression analysis was not used to determine whether either of these factors remained significant after accounting for BMC/%FFM. The authors hypothesized that DXA errors due to tissue thickness may have been one reason behind the observed association between the differences between methods and the anthropometric measurements. Their hypothesis, however, is not supported by studies that indicate that DXA overestimates (rather than underestimates) %BF at higher tissue thicknesses (73). Nonetheless, other potential contributors to variations in DXA should be considered, as discussed below.

Limitations of DXA

Because DXA does not rely on the assumptions of a 2C model to provide estimates of body composition and because it does not depend on subject performance, DXA is sometimes regarded as a standard against which other methods can be validated. However, like most other methods for measuring body composition, DXA is also subject to errors (74–76). Compared with chemical analysis, Jebb et al (75) reported that DXA underestimated the fat mass of deboned pork shoulders by 5–8% on average, whereas others reported that DXA overestimated %BF in small animals by an average of ≈30% (77, 78). Furthermore, %BF, fat mass, fat-free mass, and bone mineral estimates have been shown to vary among brands (79–82), test modes [eg, pencil beam compared with fan beam (Hologic, Waltham, MA)] (83), and software versions (84), and by tissue thickness (75, 85). Although the studies comparing the BOD POD and DXA varied in each of these respects (Table 3), no particular aspect of DXA can be singled out as a likely candidate for the lack of agreement among these studies. However, the different machines, software brands, modes, and subject thicknesses certainly contributed to the variability in the findings.

Subject sex

Of the 4 studies that compared the BOD POD and DXA measurements of %BF in men and women (20, 31, 32, 68), only the study by Koda et al (68) reported a significant effect of sex on the difference between the 2 methods. One possibility for the discrepant findings among studies is that the influence of sex on differences between these methods exists only in older subjects because Koda et al was one of only 2 studies that included older subjects. Although Nuñez et al (20) also studied older subjects, the inclusion of younger subjects as well in their study may have masked any potential effect of age in the older subjects. To better understand whether potential differences between the 2 methods are sex specific, we performed Bland-Altman analysis on group means for studies that reported mean values separately for males and females (19, 20, 29, 32, 41, 43, 68). These data are shown in **Figure 2**. There was an overall negative bias of −1.0% BF ($P = 0.10$) and no trend for differences in %BF between the BOD POD and DXA to vary by sex or with increasing %BF. The underlying reasons for the upward trend shown in Figure 1 (BOD POD compared with HW) but not in Figure 2 (BOD POD compared with DXA) should be addressed in future studies.

VALIDITY OF THE BOD POD RELATIVE TO MULTICOMPARTMENT MODELS

It is thought that the most accurate body-composition measurement, short of direct carcass analysis, can be obtained with the use of multicompartment models (86, 87). Multicompartment models are believed to give more accurate results than do the more traditional 2C models because they avoid assumptions about the density of the fat-free mass. With multicompartment models, the multiple compartments of the fat-free mass (mineral, bone, protein, and water) are actually measured, allowing for calculation of the density of fat-free mass, and the precision with which body composition can be estimated is increased (38, 62, 88, 89). Because of these advantages, the 4C model has been recommended as the new

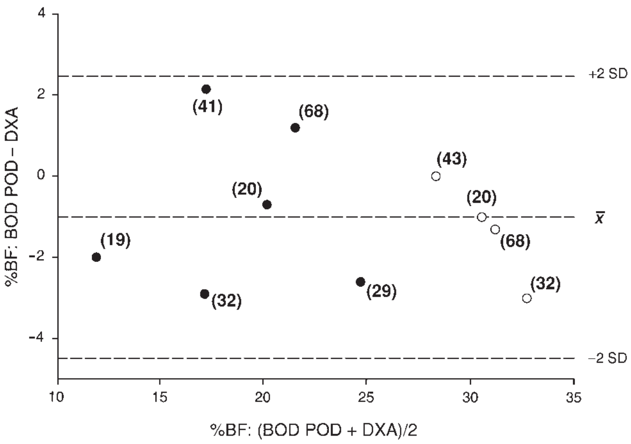


FIGURE 2. Bland-Altman plot of sex-specific mean differences between percentage body fat (%BF) measured with the BOD POD (Life Measurement, Inc, Concord, CA) and with dual-energy X-ray absorptiometry (DXA) in men (●) and women (○) in individual studies (reference numbers in parentheses). Only studies published from December 1995 to August 2001 that provided mean data for men and women separately were used. The relation between the difference between the 2 methods and the average of the 2 methods was not significant.

gold standard against which other techniques should be validated (47, 87). Three studies in adults (19, 42, 43) and one study in children (44) used a multicompartment model to validate the BOD POD. These studies are discussed below.

Summary of findings in adults

The BOD POD compared with the 4C model

Fields et al (43) studied young to middle-aged women with the use of the 4C model of Baumgartner et al (87) as the standard against which to compare the BOD POD. In this model, D_b was assessed with the BOD POD, TBW by isotopic dilution, and the bone mineral content by DXA. %BF from the BOD POD was calculated by using the 2C model of Siri (24). Although the mean difference between methods was significant (BOD POD − 4C model = −2.2% BF), the R^2 value was high (0.95) and the SEE of 2.3% BF was excellent (47). Furthermore, the 95% CI around the mean difference was relatively narrow in comparison with the wider CIs found when the BOD POD was compared with either HW or DXA in other studies, as summarized in Tables 2 and 3, ranging from −6.8% to 2.2% BF. Also of interest, the BOD POD and HW performed similarly when both were evaluated against a 4C model. As in other studies that compared the 2C densitometric model obtained from HW with a 4C model (87, 90–92), the study by Fields et al (43) found that the aqueous and mineral fractions of the fat-free mass were positively and negatively associated, respectively, with the difference in %BF calculated as BOD POD − 4C model.

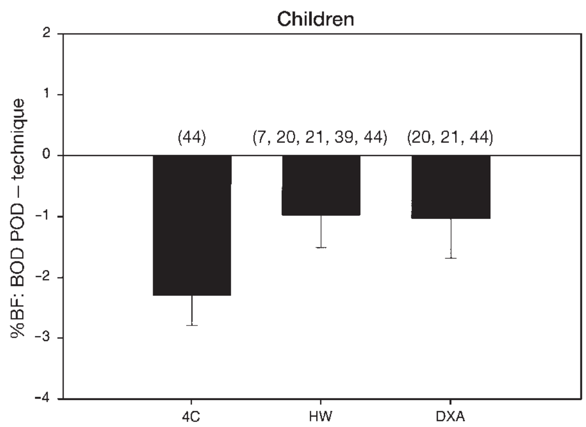
More recently, Millard-Stafford et al (42) assessed %BF with the BOD POD and HW with Siri's (24) 2C model and a 4C model (93) in 50 young men and women of mixed ethnicity (35 white, 15 black). Calculations of %BF with the 4C model were determined with the use of D_b derived from the BOD POD and from HW. %BF determined with the BOD POD differed significantly from that determined with HW when the 2C model was used, and both values differed significantly from their respective 4C models. That is, the results with the BOD POD 2C model differed significantly from those with the BOD POD 4C model, and the results from the HW 2C model differed significantly from those with the HW 4C model. The highest %BF was found with the HW 4C model (19.3% BF), followed by the BOD POD 4C and HW 2C models (each 17.8% BF) and the BOD POD 2C model (15.0% BF). Both the aqueous fraction of the body and D_b were positive predictors of the difference between the BOD POD and 4C model %BF measurements (calculated as BOD POD – 4C), and the mineral fraction of the body was a negative predictor. Limits of agreement in Bland-Altman analysis were −6.1% to 3.1% BF for individual differences between measurements made with the BOD POD 4C and HW 4C models; females tended to have positive differences and males tended to have negative differences. Nevertheless, whether subject sex per se was a significant predictor of this difference independent of %BF was not ascertained because of the small proportion of females studied and the minimal overlap in %BF between the sexes. Any potential influence of ethnicity also was not reported.

The BOD POD compared with the 3C model

Collins et al (19) compared %BF measured with the BOD POD [using the Siri (24) equation] with that calculated with a 3C density-mineral model in a subset (n = 20) of their original 69 subjects in whom the BOD POD was compared with HW (Table 2). The 3C density-mineral model was originally proposed by Lohman (47) in 1992 and later modified by Modlesky et al (94) in 1996. In this case, %BF was calculated with the use of body mineral (derived from the bone mineral content measured by DXA) and D_b from HW. Although the group mean difference was small (a difference of −1.8% BF between the BOD POD and the 3C model) and the SEE from the regression analysis was excellent (2.4% BF) per Lohman (47), the regression equation showed poor agreement between the BOD POD and the 3C model (slope = 0.65, R^2 = 0.64). No Bland-Altman analyses were presented. One reason for these relatively poor results may be that D_b derived from HW was used in the 3C model rather than D_b derived from the BOD POD. Although the advantage of this is that it allows an independent assessment of the BOD POD and the 3C model, it may have confounded the comparison because BF measurements were 2.4% lower (and thus D_b was higher) with the BOD POD than with HW in this subgroup. (As discussed earlier, this difference between the 2 methods may have been influenced by several factors.). It is also important to note that DXA measurements were not evaluated against the 3C model; therefore, it is not known whether the BOD POD performed better or worse than DXA in this population when evaluated in comparison with the 3C model.

Summary of findings in children

In the only study published thus far in which the 4C model was used in children (aged 9–14 y), Fields and Goran (44) evaluated the BOD POD and other methods. They used the 4C model of Lohman (46), which incorporates D_b derived with the BOD POD, TBW measured by isotopic dilution, and bone mineral measured with DXA. The age-adjusted 2C models of Lohman (46) were used to calculate %BF from D_b measured with the BOD POD, which was compared with %BF calculated with the use of the 4C model. Although the R^2 value was relatively high (0.90) and the SEE low (3.2% BF), the BOD POD significantly underestimated %BF; there was a difference in %BF of −2.7% between the 2 methods. However, HW also underestimated %BF by an even larger amount (difference of −3.9% BF). Additionally,



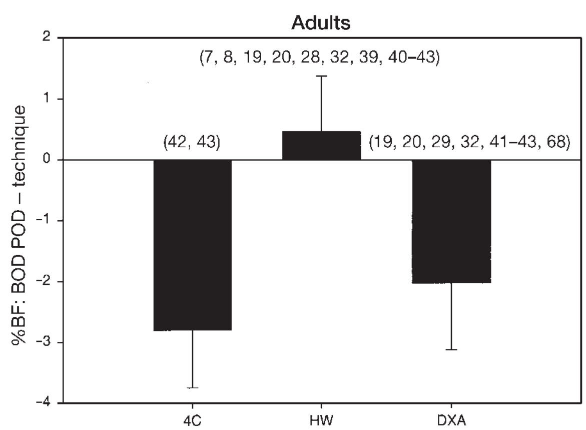


FIGURE 3. Mean (±SEM) differences between percentage body fat (%BF) measured with the BOD POD (Life Measurement, Inc, Concord, CA) and with 4-compartment (4C) models, hydrostatic weighing (HW), and dual-energy X-ray absorptiometry (DXA). Only studies published from December 1995 to August 2001 that provided mean data for men and women separately were used. Reference numbers in parentheses.

in other analyses of the 4 methods studied (BOD POD, HW, DXA, and TBW), including residual plot examination, the BOD POD was the only method that showed no significant tendency to underestimate %BF at a lower fatness and to underestimate %BF at higher fatness. Thus, in these children, the BOD POD emerged as the single best method to evaluate %BF in comparison with the gold standard estimate provided by the 4C model.

CRITICAL EVALUATION OF PREVIOUS STUDIES AND SUGGESTIONS FOR FUTURE RESEARCH

As shown in **Figure 3**, the average mean differences in %BF between the BOD POD and HW were <1% in adults and children, whereas the differences in %BF between the BOD POD and DXA were <1% in adults and 2% in children. However, it is important to note that the latter difference was based on the results of only 3 studies, the findings of which varied considerably. Taken together, the studies summarized in Tables 2 and 3 show that on average the methods agreed quite well, but there were large variations among study means. Also, the data in these tables show that there were wide limits of agreement between the methods, indicating that differences between methods for individuals can be quite large. These individual differences are attributable to both the combined imprecision of the 2 methods being compared and to disagreement between the methods.

Compared with 4C models, based on a few studies (42, 43, 44), the BOD POD underestimates %BF by ≈2–3 % in adults and children (Figure 3); a recent study (42) showed that HW underestimates %BF by a similar amount. Therefore, differences between the BOD POD and 4C model are partly explained by limitations in the assumption of the 2C models rather than to limitations in the BOD POD per se. Further support for this idea comes from several studies, which showed that variations between both the BOD POD and HW 2C models and respective 4C models are associated with deviations from the assumed chemical composition of the body (42, 43, 87, 90, 91). Errors in the 2C model are also partly responsible for the observed within-subject differences between the BOD POD and DXA, but errors in DXA itself, as discussed previously, are also responsible.

Other than the limitations of the 2C model, reasons for the differences among individuals within a study and the discrepancies among study means remain largely unknown, as illustrated by within-subject comparisons between the BOD POD and HW. Because both of these methods are based on a 2C model, they are subject to the same errors when converting D_b to %BF if the same 2C model is used for each conversion. Differences in results among studies and individuals are attributable to several factors, including differences in laboratory equipment, study design, subject characteristics, and in some cases a failure to follow the manufacturer's recommended protocol. To a large extent, the individual differences remain unexplained and future studies should be aimed at explaining these differences.

Other goals for future research include a comparison of several reference methods within a single study with the best available gold standard (either a multicomponent model or chemical analysis). This was done in one study (44), but other studies have largely focused on rigorous comparisons of the BOD POD with a reference method but no such rigorous comparison of the other methods with a reference method in the same subjects. In addition, whether there is a systematic effect of sex, independent of %BF, on differences between methods should be explored by matching men and women for %BF and for other factors that could influence differences between methods. Other potential contributors to differences between methods include errors in HW and DXA, and studies of the BOD POD should include investigation of these errors. A multilaboratory validation study would help determine whether some of the differences among the study findings can be attributed to differences between laboratory equipment. Studies of different population groups—including children, older adults, and obese subjects—and systematic investigations of differences among methods by ethnicity are also needed. Additionally, information on the reliability and validity of V_{TG} measurements, the validity of V_{TG} prediction, and factors affecting V_{TG} are necessary as are ways to improve V_{TG} measurements.

Because the BOD POD is designed to measure body volume, investigators are encouraged in future studies to include data on body volume in addition to %BF (or D_b). The software versions used with all equipment should be reported, including the BOD POD and other computerized equipment such as V_R measuring devices and DXA. Regarding study design, ≥30 subjects should be included in the studies (more if different population groups are

TABLE 4
Subjective ratings of various aspects of body-composition measurements with the BOD POD compared with reference methods[1]

	BOD POD[2]	Multicompartment models	HW	DXA	TBW
Cost	4	1	4–5	3	2–5
Time required to perform a single measurement	5	1	3	4–5	2
Equipment maintenance	4	1	2	5	1
Subject friendliness	5	1	2	5	4
User friendliness	5	2	3	3	3
Ability to accommodate a wide range of subject types[3]	4	3	1	3	5
Subject safety	5	3	4	3	5

[1] The ratings are based on a scoring system of 1 to 5, where 1 represents "least favorable" and 5 represents "most favorable." HW, hydrostatic weighing, including measurement of residual lung volume; DXA, dual-energy X-ray absorptiometry (fan or pencil beam); TBW, total body water measured by isotope dilution, including analysis by mass spectrometry or infrared spectrophotometry.

[2] Life Measurement Inc, Concord, CA.

[3] Children and subjects who are extremely obese, very tall, elderly, pregnant, or disabled, or who have musculoskeletal limitations.

compared) and, at a minimum, studies should report mean differences, regression analyses with body composition measured with the BOD POD as the independent variable and that by the reference method as the dependent variable (including goodness of fit with the line of identity, R^2, and SEE), and Bland-Altman analyses. Finally, strict adherence to the standard test conditions is imperative.

PRACTICAL ISSUES

The authors' subjective ratings of some of the practical aspects of the BOD POD in comparison with the reference methods (multicompartment models, HW, DXA, and TBW by isotope dilution) are shown in **Table 4**. Specific areas considered were cost, time required to perform a single measurement, equipment maintenance, subject and user friendliness, ability to accommodate a wide range of subject types, and subject safety. The BOD POD rated at or near the top in each category.

CONCLUSIONS

In conclusion, the BOD POD is a reliable and valid technique that can quickly and safely evaluate body composition in a wide range of subject types, including those who are often difficult to measure, such as the elderly, children, and obese individuals. More studies using multicompartment models as a reference standard are needed to validate the BOD POD for use in these and other populations. Additionally, some sources of variation between the BOD POD and other reference methods remain unknown and should be systematically studied.

We thank Paul Molé for thoughtful discussions and Sai Krupa Das, Manjiang Yao, and Paul Fuss for helpful editorial advice. None of the authors have any financial ties with any of the manufacturers of the products mentioned in this study.

REFERENCES

1. Dempster P, Aitkens S. A new air displacement method for the determination of human body composition. Med Sci Sports Exerc 1995;27:1692–7.
2. Sheng HP, Adolph AL, Smith E, Garza C. Body volume and fat-free mass determinations by acoustic plethysmography. Pediatr Res 1988; 24:85–9.
3. Dell RB. Comparison of densitometric methods applicable to infants and small children for studying body composition. Report of the 98th Ross Conference in Pediatric Research. Columbus, OH: Ross Laboratories, 1989:22–30.
4. Fomon SJ, Jensen RL, Owen GM. Determination of body volume of infants by a method of helium displacement. Ann N Y Acad Sci 1963;110:80–90.
5. Gnaedinger RH, Reineke EP, Pearson AM, Van Huss WD, Wessel JA, Montoye HJ. Determination of body density by air displacement, helium dilution, and underwater weighing. Ann N Y Acad Sci 1963;110:96–108.
6. Pearson WR. A photogrammetric technique for the estimation of surface area and volume. Ann N Y Acad Sci 1963;110: 109–12.
7. Wells JCK, Douros I, Fuller NJ, Elia M, Dekker L. Assessment of body volume using three-dimensional photonic scanning. Ann N Y Acad Sci 2000;904:247–54.
8. Iwaoka H, Yokoyama T, Nakayama T, et al. Determination of percent body fat by the newly developed sulfur hexafluoride dilution method and air displacement plethysmography. J Nutr Sci Vitaminol (Tokyo) 1998;44:561–8.
9. Petty DH, Iwanski R, Gap CX, Dressendorfer RH. Total body plethysmography for body volume determination. IEEE Frontiers Eng Computing Health Care, 1984;6:316–9.
10. Taylor A, Scopes JW, du Mont G, Taylor BA. Development of an air displacement method for whole body volume measurement of infants. J Biomed Eng 1985;7:9–17.
11. Murlin JR, Hoobler BR. The energy metabolism of normal and marasmic children with special reference to the specific gravity of the child's body. Proc Soc Exp Biol Med 1913;11:115–6.
12. Pfaundler M. Körpermass-studien an kindern. IV. Von körpervolumen und der körperdichte. (Body mass studies in children. IV. Concerning body volume and body density.) Ztschr f Kinderheilk 1916; 14:123–37 (in German).
13. Friis-Hansen B. The body density of newborn infants. Acta Paediatr 1963;52:513–21.
14. Gundlach BL, Visscher GJW. The plethysmometric measurement of total body volume. Hum Biol 1986;58:783–99.
15. Life Measurement, Inc. BOD POD body composition system: operator's manual. Concord, CA: Life Measurement, Inc, 1997.
16. Daniels F, Alberty RA. Physical chemistry. New York: John Wiley and Sons, Inc, 1967.
17. DuBois D, DuBois EF. A formula to estimate the approximate surface area if height and weight be known. Arch Intern Med 1916; 17:863–71.
18. DuBois AB, Botelho ST, Bedell GN, Marshall R, Comroe JH. A rapid plethysmographic method for measuring thoracic gas volume: a comparison with a nitrogen washout method for measuring functional residual capacity in normal subjects. J Clin Invest 1956;35: 322–6.
19. Collins MA, Millard-Stafford ML, Sparling PB, et al. Evaluation of the BOD POD for assessing body fat in collegiate football players. Med Sci Sports Exerc 1999;31:1350–6.

20. Nuñez C, Kovera AJ, Pietrobelli A, et al. Body composition in children and adults by air displacement plethysmography. Eur J Clin Nutr 1999;53:382–7.
21. Lockner DW, Heyward VH, Baumgartner RN, Jenkins KA. Comparison of air-displacement plethysmography, hydrodensitometry, and dual X-ray absorptiometry for assessing body composition of children 10 to 18 years of age. Ann N Y Acad Sci 2000;904:72–8.
22. Crapo RO, Morris AH, Clayton PD, Nixon CR. Lung volumes in healthy nonsmoking adults. Bull Eur Physiopathol Respir 1982;18: 419–25.
23. Behnke AR, Feen BG, Welham WC. The specific gravity of healthy men. JAMA 1942;118:495–8.
24. Siri WE. Body composition from fluid spaces and density: analysis of methods. In: Brozek J, Henschel A, eds. Techniques for measuring body composition. Washington, DC: National Academy of Sciences, National Research Council, 1961:223–4.
25. Brozek J, Grande F, Anderson JT, Keys A. Densitometric analysis of body composition: revision of some quantitative assumptions. Ann N Y Acad Sci 1963;110:113–40.
26. Schutte JE, Townsend EJ, Hugg J, Shoup RF, Malina RM, Blomqvist CG. Density of lean body mass is greater in blacks than in whites. J Appl Physiol 1984;56:1647–9.
27. Wagner DR, Heyward VH. Validity of two-compartment models for estimating body fat of black men. J Appl Physiol 2001;90:649–56.
28. McCrory MA, Gomez TD, Bernauer EM, Molé PA. Evaluation of a new air displacement plethysmograph for measuring human body composition. Med Sci Sports Exerc 1995;27:1686–91.
29. Sardinha LB, Lohman TG, Teixeira P, Guedes DP, Going SB. Comparison of air displacement plethysmography with dual-energy X-ray absorptiometry and 3 field methods for estimating body composition in middle-aged men. Am J Clin Nutr 1998;68:786–93.
30. Biaggi RR, Vollman MW, Nies MA, et al. Comparison of air-displacement plethysmography with hydrostatic weighing and bioelectrical impedance analysis for the assessment of body composition in healthy adults. Am J Clin Nutr 1999;69:898–903.
31. Miyatake N, Nonaka K, Fujii M. A new air displacement plethysmograph for the determination of Japanese body composition. Diabetes Obes Metab 1999;1:347–51.
32. Levenhagen DK, Borel MJ, Welch DC, et al. A comparison of air displacement plethysmography with three other techniques to determine body fat in healthy adults. JPEN J Parenter Enteral Nutr 1999; 23:293–9.
33. Pierson RN Jr, Wang J, Heymsfield SB, et al. Measuring body fat: calibrating the rulers. Intermethod comparisons in 389 normal Caucasian subjects. Am J Physiol 1991;261:E103–8.
34. Van Der Ploeg GE, Gunn SM, Withers RT, Modra AC, Crockett AJ. Comparison of two hydrodensitometric methods for estimating percent body fat. J Appl Physiol 2000;88:1175–80.
35. Lohman TG. Dual energy X-ray absorptiometry. In: Roche AF, Heymsfield SB, Lohman TG, eds. Human body composition. Champaign, IL: Human Kinetics Publishers, 1996.
36. Economos CD, Nelson ME, Fiatarone MA, et al. A multi-center comparison of dual energy X-ray absorptiometers: in vivo and in vitro soft tissue measurement. Eur J Clin Nutr 1997;51:312–7.
37. Santana H, Zoico E, Turcato E, et al. Relation between body composition, fat distribution and lung function in elderly men. Am J Clin Nutr 2001;73:827–31.
38. Wells JCK, Fuller NJ. Precision of measurement and body size in whole-body air-displacement plethysmography. Int J Obes Relat Metab Disord 2001;25:1161–7.
39. Dewit O, Fuller NJ, Fewtrell MS, Elia M, Wells JCK. Whole body air displacement plethysmography compared with hydrodensitometry for body composition analysis. Arch Dis Child 2000;82:159–64.
40. Fields DA, Hunter GR, Goran MI. Validation of the BOD POD with hydrostatic weighing: influence of body clothing. Int J Obes Relat Metab Disord 2000;24:200–5.
41. Wagner DR, Heyward VH, Gibson AL. Validation of air displacement plethysmography for assessing body composition. Med Sci Sports Exerc 2000;32:1339–44.
42. Millard-Stafford ML, Collins MA, Evans EM, Snow TK, Cureton KJ, Rosskopf LB. Use of air displacement plethysmography for estimating body fat in a four-component model. Med Sci Sports Exerc 2001;33:1311–7.
43. Fields DA, Wilson GD, Gladden LB, Hunter GR, Pascoe DD, Goran MI. Comparison of the BOD POD with the four-compartment model in adult females. Med Sci Sports Exerc 2001;33:1605–10.
44. Fields DA, Goran MI. Body composition techniques and the four-compartment model in children. J Appl Physiol 2000;89:613–20.
45. Wells JC, Fuller NJ, Dewit O, Fewtrell MS, Elia M, Cole TJ. Four-compartment model of body composition in children: density and hydration of fat-free mass and comparison with simpler models. Am J Clin Nutr 1999;69:904–12.
46. Lohman TG. Assessment of body composition in children. Pediatr Exerc Sci 1989;1:19–30.
47. Lohman TG. Advances in body composition assessment. Champaign, IL: Human Kinetics Publishers, 1992.
48. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;8: 307–10.
49. McCrory MA, Molé PA, Gomez TD, Dewey KG, Bernauer EM. Body composition by air-displacement plethysmography by using predicted and measured thoracic gas volumes. J Appl Physiol 1998; 84:1475–9.
50. Rosenthal M, Cramer D, Bain SH, Denison D, Bush A, Warner JO. Lung function in white children aged 4 to 19 years: II. Simple breath analysis and plethysmography. Thorax 1993;48:803–8.
51. Zapletal A, Paul T, Samanek M. Normal values of static pulmonary volumes and ventilation in children and adolescents. Cesk Pediatr 1976;31:532–9.
52. Higgins PB, Fields DA, Gower BA, Hunter GR. The effect of scalp and facial hair on body fat estimates by the BOD POD. Obes Res 2001;9:326–30.
53. Bedell GN, Marshall R, Dubois AB, Harris JH. Measurement of the volume of gas in the gastrointestinal tract. Values in normal subjects and ambulatory patients. J Clin Invest 1956;35:336–45.
54. Hanson CF, Winterfeldt EA. Dietary fiber effects on passage rate and breath hydrogen. Am J Clin Nutr 1985;42:44–8.
55. Levitt MD, Hirsch P, Fetzer CA, Sheahan M, Levine AS. H_2 excretion after ingestion of complex carbohydrates. Gastroenterology 1987;92:383–9.
56. Brown R, Hoppin FG, Ingram RH Jr, Saunders NA, McFadden ER Jr. Influence of abdominal gas on the Boyle's law determination of thoracic gas volume. J Appl Physiol 1978;44:469–73.
57. Habib MP, Engel LA. Influence of panting technique on the plethysmographic measurement of thoracic gas volume. Am Rev Resp Dis 1978;117:265–71.
58. McCrory MA, Fuss PJ, Saltzman E, Hays NP, Roberts SB. Body composition measurements by air-displacement plethysmography and underwater weighing: effects of gas-producing and gas-containing foods. FASEB J 2000;14:A498 (abstr).
59. Buskirk ER. Underwater weighing and body density: a review of procedures. In: Brozek J, Henschel A, eds. Techniques for measuring body composition. Washington, DC: National Academy of Sciences, National Research Council, 1961:90–105.
60. Akers R, Buskirk ER. An underwater weighing system utilizing "force cube" transducers. J Appl Physiol 1969;26:649–52.
61. Forsyth R, Plyley MJ, Shephard RJ. Residual volume as a tool in body fat prediction. Ann Nutr Metab 1988;32:62–7.
62. Friedl KE, DeLuca JP, Marchitelli LJ, Vogel JA. Reliability of body-fat estimations from a four-compartment model by using density, body water, and bone mineral measurements. Am J Clin Nutr 1992; 55:764–70.

63. Dahlback GO, Lundgren C. Pulmonary air-trapping induced by water immersion. Aerospace Med 1972;43:768–74.
64. Robertson CH Jr, Engle CM, Bradley ME. Lung volumes in man immersed to the neck: dilution and plethysmographic techniques. J Appl Physiol 1978;44:679–82.
65. Demedts M, van de Woestijne KP. Which technique for total lung capacity measurement? Bull Eur Physiopathol Respir 1980;16:705–9.
66. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. Eur Respir J Suppl 1993;16:5–40.
67. Kendrick AH. Comparison of methods of measuring static lung volume. Monaldi Arch Chest Dis 1996;51:431–9.
68. Koda M, Tsuzuku S, Ando F, Niino N, Shimokata H. Body composition by air displacement plethysmography in middle-aged and elderly Japanese: comparison with dual-energy X-ray absorptiometry. Ann N Y Acad Sci 2000;904:484–8.
69. Roemmich JN, Clark PA, Weltman A, Rogol AD. Alterations in growth and body composition during puberty. I. Comparing multicompartment body composition models. J Appl Physiol 1997;83:927–35.
70. Wang J, Heymsfield SB, Aulet M, Thornton JC, Pierson RN Jr. Body fat from body density: underwater weighing vs. dual-photon absorptiometry. Am J Physiol 1989;256:E829–34.
71. Tobe H, Tanaka S, Koda M, Satake T, Hosoi T, Orimo H. Effects of bone mineral content and density on accuracy of body fat measurement by underwater weighing. Jpn J Phys Fitness Sports Med 1996;45:503–10.
72. Kohrt WM. Preliminary evidence that DEXA provides an accurate assessment of body composition. J Appl Physiol 1998;84:372–7.
73. Jebb SA. Measurement of soft tissue composition by dual energy X-ray absorptiometry. Br J Nutr 1997;77:151–63.
74. Roubenoff R, Kehayias JJ, Dawson-Hughes B, Heymsfield SB. Use of dual-energy x-ray absorptiometry in body composition studies: not yet a “gold standard”. Am J Clin Nutr 1993;58:589–91.
75. Jebb SA, Goldberg GR, Jennings G, Elia M. Dual-energy X-ray absorptiometry measurements of body composition: effects of depth and tissue thickness, including comparisons with direct analysis. Clin Sci (Colch) 1995;88:319–24.
76. Van Loan MD. Is dual-energy X-ray absorptiometry ready for prime time in the clinical evaluation of body composition? Am J Clin Nutr 1998;68:1155–6.
77. Brunton JA, Weiler HA, Atkinson SA. Improvement in the accuracy of dual energy X-ray absorptiometry for whole body and regional analysis of body composition: validation using piglets and methodologic considerations in infants. Pediatr Res 1997;41:590–6.
78. Rose BS, Flatt WP, Martin RJ, Lewis RD. Whole body composition of rats determined by dual energy X-ray absorptiometry is correlated with chemical analysis. J Nutr 1998;128:246–50.
79. Horber FF, Thomi F, Casez JP, Fonteille J, Jaeger P. Impact of hydration status on body composition as measured by dual energy X-ray absorptiometry in normal volunteers and patients on hemodialysis. Br J Radiol 1992;65:895–900.
80. Tothill P, Avenell A, Reid DM. Precision and accuracy of measurements of whole-body bone mineral: comparisons between Hologic, Lunar and Norland dual-energy X-ray absorptiometers. Br J Radiol 1994;67:1210–7.
81. Modlesky CM, Lewis RD, Yetman KA, Rose B, Rosskopf LB, Sparling PB. Comparison of body composition and bone mineral measurements from two DXA instruments in young men. Am J Clin Nutr 1996;64:669–76.
82. Kistorp CN, Svednsen OL. Body composition analysis by dual energy X-ray in female diabetics differ between manufacturers. Eur J Clin Nutr 1997;51:449–54.
83. Clasey JL, Hartmin ML, Kanaley JA, et al. Body composition by DEXA in older adults: accuracy and influence of scan mode. Med Sci Sports Exerc 1997;29:560–7.
84. Van Loan MD, Keim NL, Berk K, Mayclin PL. Evaluation of body composition by dual energy x-ray absorptiometry and two different software packages. Med Sci Sports Exerc 1995;27:587–91.
85. Laskey MA, Lyttle KD, Flaxman ME, Barber RW. The influence of tissue depth and composition on the performance of the Lunar dual-energy X-ray absorptiometer whole-body scanning mode. Eur J Clin Nutr 1992;46:39–45.
86. Heymsfield SB, Wang J, Lichtman S, Kamen Y, Kehayias J, Pierson RN Jr. Body composition in elderly subjects: a critical appraisal of clinical methodology. Am J Clin Nutr 1989;50(suppl):1167S–75S.
87. Baumgartner RN, Heymsfield SB, Lichtman S, Wang J, Pierson RN Jr. Body composition in elderly people: effect of criterion estimates on predictive equations. Am J Clin Nutr 1991;53:1345–53.
88. Fuller NJ, Jebb SA, Laskey MA, Coward WA, Elia M. Four-component model for the assessment of body composition in humans: comparison with alternative methods, and evaluation of the density and hydration of fat-free mass. Clin Sci 1992;82:687–93.
89. Goran MI, Toth MJ, Poehlman ET. Assessment of research-based body composition techniques in healthy elderly men and women using the 4-compartment model as a criterion method. Int J Obes Relat Metab Disord 1998;22:135–42.
90. Hewitt MJ, Going SB, Williams DP, Lohman TG. Hydration of the fat-free body mass in children and adults: implications for body composition assessment. Am J Physiol 1993;265:E88–95.
91. Bergsma-Kadijk JA, Baumeister B, Deurenberg P. Measurement of body fat in young and elderly women: comparison between a four-component model and widely used reference methods. Br J Nutr 1996;75:649–57.
92. Withers RT, LaForgia J, Pillans RK, et al. Comparisons of two-, three-, and four-compartment models of body composition analysis in men and women. J Appl Physiol 1998;85:238–45.
93. Lohman TG. Applicability of body composition techniques and constants for children and youths. Exerc Sport Sci Rev 1986;14:325–57.
94. Modlesky CM, Cureton KJ, Lewis RD, Prior BM, Sloninger MA, Rowe DA. Density of the fat-free mass and estimates of body composition in male weight trainers. J Appl Physiol 1996;80:2085–96.