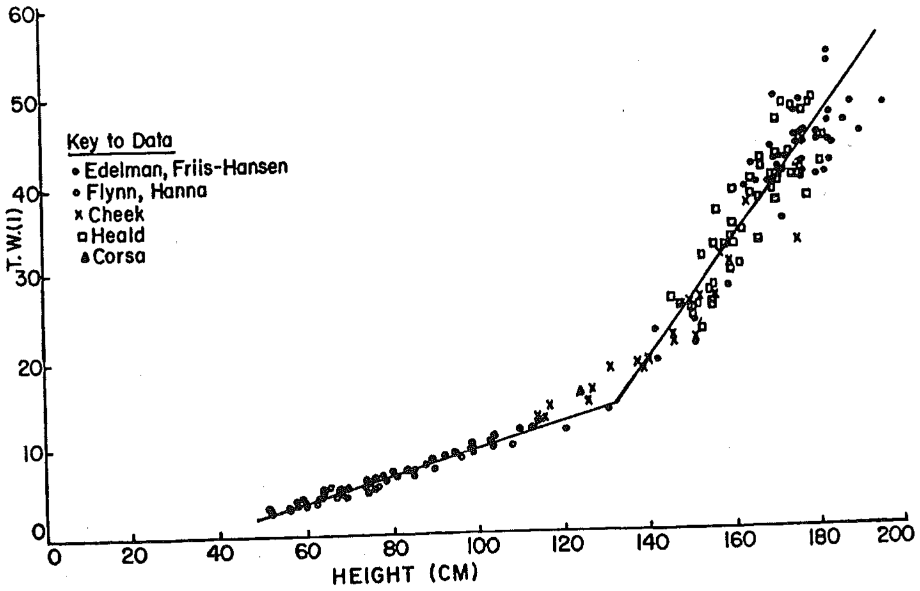

Figure 10–2. Total body water in relation to height. (From Mellits and Cheek, 1970; reproduced with permission).

Measurement of active cell mass is usually based on the determination of total body potassium, using the fact that body potassium is largely intracellular. This measurement is best accomplished by counting the gamma-ray emissions of potassium-40, a naturally occurring radioisotope. Because this method measures a natural isotope (which occurs as a known proportion of all potassium isotopes), no administration of a tracer is required; however, the method requires a special, generally unavailable, counting chamber (Lukaski, 1987; Jebb and Elia, 1993).

Bone mass is the other major component of lean body mass. Bone mass measurement can enhance the accuracy of other body components that are measured indirectly and is of inherent interest in the study of fractures. Photon absorptiometry is most commonly used to measure bone mass and is based on the principle that the mineral content of the bone being studied is directly proportional to the absorbed energy from a photon beam emitted by a radionuclide. Lukaski (1987) provides a detailed review of this technology and its accuracy. Single photon absorptiometry uses a simple and relatively inexpensive device to measure bone mass or density in the arm or legs; a site over the distal radius is the most commonly used. Although highly reproducible, even multiple measurements do not provide an accurate assessment of total bone mass, in part because much of the bone mass is contained in the axial skeleton and the densities at various sites are only moderately correlated. Dual photon absorptiometry uses the differential absorption of photon beams with two distinct energy levels to differentiate between bone mineral and soft tissue mass. Coupled with a body scanning device, this provides an accurate measurement of total body bone mass. A similar method, dual-energy x-ray absorptiometry (DEXA), in which an x-ray

generator replaces the photon source, has more recently been used to measure bone mass as well as lean and fat mass (see below).

MEASUREMENT OF RELATIVE BODY COMPOSITION

Because of widespread interest in the health effects of obesity, the most frequent use of anthropometry in epidemiologic studies is to estimate adiposity. In general, adiposity has been expressed as percent body fat (fat mass/total mass × 100%), although there may be reason to question whether this is truly the most biologically relevant measure of adiposity (see below). In epidemiologic studies, the most commonly used methods to estimate relative body composition are combinations of weight and height, skinfold thickness, and body circumferences. Newer methods based on electrical resistance and impedance, DEXA, magnetic resonance imaging, and computer-assisted tomography have become available. Because these methods measure adiposity indirectly, it is crucial to consider the degree of error associated with their use, as well as their feasibility. Because densitometry has been the generally accepted standard for measuring the percentage of body weight that is fat, this method is described first, even though it is impractical for most epidemiologic applications.

Densitometry

Densitometry (also called hydrostatic weighing) is based on the principle that fat tissue is lighter than fat-free tissue. The ratio of weights measured in air and under water, therefore, provides an estimate of the proportion of total body mass that is composed of fat. In the most widely used technique, subjects wearing a swimming suit are submerged seated on a scale in a tank of water (with a known weight strapped to their body so that they do not float). Because the air in the lungs influences weight under water, residual lung volume is measured by having subjects breath through a snorkel into a special device (for a detailed description of densitometry, see Going, 1996). Formulas for the calculation of percentage of body fat from these data have been developed by Siri (1961) and Brozek and colleagues (1963). Both biologic variation in the density of fat and lean body mass and technical variation in the measurement of density contribute to error in estimating body fat composition by densitometry. Of these factors, variation in the water content of the lean body mass, in bone size, and the density of bone appear to be the primary sources of error and may lead to errors of 3% to 4% in predicting body fatness (Lohman, 1981; Lukaski, 1987). Within a population of similar age, sex, and race, however, the biologic sources of error should be considerably less important than in the heterogeneous groups in which sources of error have been evaluated. For demographically homogeneous groups, the magnitude of error associated with densitometry is not well defined.

Combinations of Weight and Height

Weight and height are the most commonly available anthropometric measurements in epidemiologic settings; the literature regarding methods of combining them to best represent adiposity is enormous. The criteria usually employed are (1) that the index should be highly correlated with percent body fat and (2) that the index be uncorrelated with height. The first criterion is obviously most important, but also the most difficult to evaluate because a perfect standard for adiposity is not available and the best methods, such as densitometry, are difficult for practical reasons. More attention has been focused on the second criterion, probably because it is far easier to evaluate. This criterion, however, has become less important with the advent of computers because many multivariate procedures are widely available that can easily provide a statistical adjustment for height (discussed later). If height is not associated with the disease being investigated, the second criterion is largely irrelevant in that setting. Moreover, the second criterion makes the implicit assumption that adiposity is unrelated to height. This appears to be generally true for adults, but is not for children; before puberty, obese children tend to be taller than lean children (discussed later and by Roche, 1984). The two most commonly employed measures of obesity are relative weight (a standardized ratio) and indices of weight and height that are not related to a standard.

Relative weight is the ratio of a subject's observed weight to a standard or expected weight; this may also be expressed as a percentage above or below the standard. The standard weights are frequently derived from a large group of persons of the same height, sex, and (sometimes) age. These may be obtained from an external population, such as the widely used Metropolitan Life Insurance "desirable weights" (Metropolitan Life Insurance Co, 1959); these are based on associations with minimal mortality among insurees and are revised periodically. In some large investigations, such as the American Cancer Society cohort of 750,000 men and women (Lew and Garfinkel, 1979), the average weights for study participants of the same height, sex, and age are used as standards.

The use of relative weight provides a readily interpretable measure; to say that a group of subjects was 140% to 150% of the average weight for their age, sex, and height conveys a meaningful image to almost any reader. The distinct disadvantage of this approach is that findings from different studies are difficult to compare as a wide variety of standards may have been employed. It is not often appreciated, for example, that the Metropolitan Life standards are substantially below the standards based on average weights in other U.S. studies (Manson et al., 1987). Differences in standards are likely to be even greater internationally.

Obesity indices are combinations of weight and height that are not related to a standard. More than 100 years ago, Quetelet (1869) pointed out that weight/height^2 (also called the body mass index or BMI) was minimally correlated with height. Other investigators have advocated the use of weight/height, weight/height$^{1.5}$, and weight/height^3. Collectively, these have been called power indices. Benn (1971) has advocated the use of an empirically fit value for the exponent of height (p) based on the specific population being studied so that, by definition, the index (weight/height^p) is uncorrelated with height. He has further shown that such an index is perfectly correlated with relative weight based on a standard from within the same population. The use of these obesity indices, with the exception of the Benn index, has the considerable advantage that they provide measurement scales that do not vary from study to study, thus facilitating the comparison of findings. The relative merits of the different indices should be considered on the basis of their relationships with true adiposity and, to a much lesser degree, on being uncorrelated with height.

Multivariate adjustment for height provides a simple alternative to the use of relative weight or obesity indices. Weight and height can both be entered as independent variables in a multiple regression model predicting the outcome of interest; this provides a measure of the effect of weight independent of height, thus by definition weight uncorrelated with height. Conceptually, this can be thought of as the effect of weight among individuals of identical height.

Although the meaning of weight in this multivariate model containing weight and height is relatively clear, the converse is not; the interpretation of height adjusted for weight is conceptually unclear and of little interest as it is strongly related to body composition. (This was appreciated by the wit of yore who commented that he was not fat, just short for his weight.) If both height and weight-adjusted-for-height (an estimate of obesity) are of interest, a two-step procedure analogous to that suggested for adjusting nutrient intake for total caloric intake can be employed (see Chapter 11). First, a simple regression model is used with height as the independent variable (x) and weight as the dependent variable (y). The resulting residual of weight on height provides, by definition, a measure of weight uncorrelated with height. (This measure thus has all the advantages of the Benn index, with the added feature of having the usual scale of weight, i.e., kilograms or pounds.) In the second step, height and the residual of weight on height can both be entered in the multivariate model. This simultaneously provides the full effect of height as well as a measure of weight independent of height.

The assumption that true adiposity is unrelated to height is central to the interpretation of relative weight and obesity indices. This relationship has been addressed by examining the correlations between height and adiposity measured by skin folds or by densitometry (Table 10–1). In adults, it does appear that the correlation between height and adiposity is minimal; in only the data of Womersley and Durnin (1977) does there appear to be a slight inverse relationship.

Among the obesity indices shown in Table 10–1, the correlation with height has generally been lowest for BMI (weight/height^2). The Benn index, using an empirically fit exponent for weight (Benn, 1971; Lee et al., 1981), does not seem to provide any clear advantage as the correlation between height and weight/height^2 is typically already small (Garn and Pesick, 1982; Colliver et al., 1983). Others have found that the exponent of 1.5 produces a slightly lower correlation with height among women (Micozzi et al., 1986).

Relationships of height and weight among children are considerably more complex, probably because this is a period of active growth; an exhaustive review is beyond the scope of this book. At some ages, however, height is positively associated with adiposity to a degree such that an assumption of independence is materially violated (Killeen et al., 1978). In this situation one could use the combination of weight and height most strongly correlated with obesity and adjust for height, if it is related to the outcome being studied, with multivariate analysis. Alternatively, the use of a more direct measure, such as skinfolds, may be preferable.

The validity of combinations of weight and height as measures of adiposity has frequently been assessed by correlating these with skinfold thickness (Table 10–2). Because skinfold thicknesses themselves are imperfect indicators of adiposity, the absolute value of these correlations should not be interpreted as a direct measure of validity. Comparing the correlations of skinfolds with different combinations of weight and height, however, may be useful to assess their relative degrees of validity. As shown in Table 10–2, the correlations with skinfolds are quite similar whether one uses weight/height or weight/height^2, and the correlations with weight/height^3 are only slightly reduced. Indeed, the use of weight alone is nearly as good as any weight index. This result is not surprising as it can be readily appreciated that most variation in weight between individuals is independent of height because adult heights do not vary dramatically; within one age and sex group the range from the tallest to the shortest is typically only about 20%.

The Influence of Frame Size

It is commonly assumed that weight should be evaluated in relation to frame size and that an accurate measurement of skeletal dimensions may improve the validity of obesity indices that are based simply on weight and height. Indeed, height itself is

Table 10–1. Correlation coefficients between height and selected indices of obesity (Pearson r)

Source	Subjects	Wt	Wt/Ht	Wt/Ht²	Wt/Ht³	Wt/$\overline{Wt}$	Log skinfold	Body fat (densitometry)
Allen et al. (1956)[a]	55 men	—	0.41	0.16	−0.16	—	—	0.03
	26 women		0.27	0.03	−0.24	—	—	0.05
Khosla and Lowe (1967)	5,000 men (15–64 yr)	0.43/0.59	0.19/0.37	−0.10/0.08	−0.36/−0.24	—	—	—
Evans and Prior (1969)	432 men	0.44	0.25	0.02	−0.21	—	−0.13[b]/0.15[b]	—
	378 women	0.55	0.33	0.05	−0.18		−0.06[b]/0.19[b]	—
Florey (1970)	1,723 men	0.48	—	—	—	—	0.12 (trcps) 0.04 (infscp) 0.02 (trcps) −0.08 (infscp)	—
	2,202 women	0.25	—	—	—	—	−0.08 (infscp)	
Keys et al. (1972)	180 students	0.46	0.25	0.02	−0.24[c]	0.04	0.04	—
	249 executives	0.40	0.18	0.06	−0.30[c]	0.01	0.02	—
Goldbourt and Medalie (1974)	9,475 men	0.52	0.28	−0.03	−0.31[c]	—	0.09 (trcps) 0.05 (infscp)	—
Womersley and Durnin (1977)	245 men	—	0.01	−0.22	−0.43[c]	−0.08	−0.13	−0.22
	324 women	—	0.06	−0.10	−0.26[c]	−0.04	−0.06	−0.13
Revicki and Israel (1986)	474 men	0.19	−0.03	−0.24	—	0.01	0.02	—
Killeen et al. (1978)	13,867 children 6–17 yr (by race and sex)	0.21/0.71	−0.01/0.38[d]	−0.36/0.18	—	—	0.02 0.34 (infscp)	—
Michielutte et al. (1984)	832 boys	—	0.76	0.44	−0.27	—	—	—
	836 girls aged 5–12 yr	—	0.78	0.50	−0.18	—	—	—
Micozzi et al. (1986)	5,808 men (25–74 yr)	0.42	0.08/0.24	~0.00/ −0.08	−0.22/−0.33	—	—	—
	8,592 women (25–74 yr)	0.21	~0.00/0.08	~0.00/ −0.13	−0.25/−0.31	—	—	—

[a] Calculations from Womersley and Durnin (1977).
[b] Sum of triceps (trcps) and infrascapular (infscp).
[c] Correlations are for cube root of weight divided by height (ponderal index).
[d] Strongest positive correlations for youngest children.

Table 10–2. Correlation of skinfold measures with anthropometric indices of obesity (Pearson r)

Source	Subjects	Skinfold	Ht	Wt	Wt/Ht	Wt/Ht²	Wt/Ht³
Flory (1970)	1,723 men	Triceps	0.12	0.44	0.45	0.42	0.36[a]
	1,723 men	Infrascapular	0.04	0.59	0.64	0.64	0.59[a]
	2,202 women	Triceps	0.02	0.47	0.47	0.46	0.44[a]
	2,202 women	Infrascapular	0.08	0.61	0.65	0.66	0.64[a]
Keys et al. (1972)	180 students 18–24 yr	Sum of triceps + infrascapular	0.06	0.78	0.83	0.85	0.81[a]
	249 executives		0.00	0.72	0.77	0.78	0.74[a]
Goldbourt and Medalie (1974)	9,475 Israeli men	Triceps	0.09	0.39	0.42	0.40	0.36[a]
		Infrascapular	0.05	0.54	0.60	0.60	0.56[a]
Killen et al. (1978)	13,687 children 6–17 yr by age and sex	Infrascapular	—	—	0.55–0.81	0.61–0.83[b]	0.47–0.81
Michielutte et al. (1984)	832 boys 5–12 yr	Triceps	—	—	0.73	0.81[b]	0.69
	835 girls 5–12 yr	Triceps	—	—	0.73	0.81[b]	0.64
Revicki and Israel (1986)	474 men	7 skinfolds (computed % fat)	0.01	0.71 0.70[c]	0.75 0.74[c]	0.76 0.72[c]	0.73 0.72[c]
Micozzi et al. (1986)	5,808 men	Infrascapular	~0.00	0.69	0.75	0.77	0.74
	8,592 women	Infrascapular	−0.09	0.76	0.79	0.80	0.79

[a] Correlations for cube root of weight divided by height (ponderal index).
[b] Correlations lowest for younger children.
[c] Age-adjusted.

basically a one-dimensional measure of frame size that is easily available in most studies. Widely used "ideal weights," such as those published by Metropolitan Life, are often provided for small, medium, and large frame sizes. These categories of frame size, however, have no quantitative definition and are left to individual judgment.

In addition to height, other measures of skeletal dimensions include biacromial diameter, knee and elbow width (Frisancho, 1984), biiliac diameter, and chest depth (Garn et al., 1986); such measurements have sometimes been combined into indices of frame size (Katch and Freedson, 1982). Katch and colleagues (1982) have demonstrated that frame sizes based on self-report or on a subjective rating by an expert correspond poorly with a standardized measurement of frame size. Roche (1984) has reviewed studies that address the issue of whether measures of frame size improve the prediction of body fat composition above and beyond that provided by simple weight and height. Overall, there appears to be no consistent evidence that frame size measurements in addition to height provide any important refinement in the estimation of obesity. This is probably expected because, as demonstrated by the data in Tables 10–2 and 10–3, even height provides only a modest incremental improvement in prediction of body fat composition; further refinements in frame size estimation are likely to produce smaller marginal gains. Although additional work is warranted to identify simple measures of frame size that may improve the interpretation of weight, the cost and difficulty involved in obtaining such measurements are unlikely to be justified in epidemiologic studies of obesity.

Table 10–3. Correlation coefficients between densitometry estimates of body fat composition and anthropometric indices of obesity (Pearson *r*)

Source	Subjects	Wt	Wt/Ht	Wt/Ht^2	Wt/Ht^3	Skinfold
Allen et at. (1956)[a]	55 men	—	0.70	0.72	0.68	—
	26 women	—	0.74	0.80	0.77	—
Parizkova (1961)	62 normal adolescents	—	—	—	—	0.74 (triceps) 0.80 (infrascapular)
Seltzer et al. (1965)	32 obese adolescent girls	—	—	—	—	0.69 (triceps) 0.59 (infrascapular)
Keys et al. (1972)	180 students	—	0.83	0.85	0.79[b]	0.85[c]
	249 executives	—	0.66	0.67	0.66[b]	0.82[c]
Womersley and Durnin (1977)	245 men	—	0.68	0.71	0.72[b]	0.84
	324 women	—	0.81	0.82	0.84[b]	0.86
Harsha et al.	242 black and white children	0.32	—	—	—	0.81 (triceps) 0.76 (infrascapular)
Roche et al. (1981)	68 boys (6–12 yr)	0.33	0.73	0.68	0.74	0.84 (triceps) 0.74 (infrascapular)
	49 girls (6–12 yr)	0.23	0.69	0.55	0.62	0.83 (triceps) 0.68 (infrascapular)
	63 boys (13–17 yr)	0.30	0.71	0.61	0.71	0.78 (infrascapular) 0.72 (infrascapular)
	81 girls (13–17 yr)	0.72	0.77	0.77	0.74	0.83 (triceps) 0.81 (infrascapular)
	141 men (18–49 yr)	0.67	0.64	0.77	0.75	0.70 (triceps) 0.75 (infrascapular)
	135 women (18–49 yr)	0.70	0.69	0.76	0.75	0.77 (triceps) 0.71 (infrascapular)
Revicki and Israel (1986)	474 men	0.66	0.70	0.71	0.69	0.84 (7 measures)
		—	0.52[d]	0.58[d]	0.58[d]	

[a] Calculations from Womersley and Durnin (1977).

[b] Correlations are for cube root of weight divided by height (ponderal index).

[c] Sum of triceps and infrascapular.

[d] Age-adjusted.

Skinfold Measurements

Next to combinations of height and weight, skinfolds are probably the most widely used method to measure body composition in epidemiologic studies. This method has conceptual appeal because it provides a direct measure of body fat; its major limitations are that not all fat is accessible to the calipers (such as intraabdominal and intramuscular fat) and that the distribution of subcutaneous fat can vary considerably over the body. This variability in distribution of subcutaneous fat creates difficulties when measurements at one or only a few sites are used to represent overall body fat composition; however, these distributions may be of interest in their own right (discussed later).

The technical aspects of skinfold measurements and considerations for selecting specific sites are discussed in detail elsewhere (Habicht et al., 1979; Lohman, 1981; Mueller and Stallones, 1981; Rose and Blackburn, 1982; Roche, 1984; Roche et al., 1996). In general, skinfold measurements are substantially less reproducible than most other anthropometric measures, such as weight, height, and limb and girth circumferences (Bray et al., 1978; Habicht et al., 1979; Lukaski, 1987).

Ruiz and colleagues (1971) formally investigated sources of variation in skinfold thickness measurements. They found that a small difference (2.5 cm) in the site of measuring the triceps skinfold, for example, resulted in a difference as large as 50% in the average skinfold. Other factors, such as the manner in which the skinfold was picked up and the depth of caliper bite, contributed less to variation. Jointly, these factors contribute to the substantial interobserver variation that has typically been reported for such measurements. Because of this relatively high degree of error variance, skinfold thickness measurements are of limited use in following changes in obesity over time (Bray et al., 1978).

The validity of skinfold thickness assessed by calipers as a measure of true subcutaneous adipose thickness (as opposed to being a measure of body fat composition, which is discussed next) has been assessed by comparing data obtained by calipers and by computed tomography or ultrasound (Roche, 1996). Fanelli and Kuczmarski (1984) and Kuczmarski and colleagues (1987) found that subcutaneous fat measurements by ultrasound were not superior to skinfold measurements by caliper in predicting body fat composition determined by densitometry among relatively lean individuals. Among obese adults, however, the ultrasonic measurements proved to be superior. Among the obese group, correlations between skinfold thickness and ultrasonic measurements at the same site ranged from 0.30 (waist) to 0.72 (thigh and biceps). Abe and colleagues (1994), however, found that subcutaneous adipose estimated by ultrasound was overall more predictive of body fat measured by densitometry than when measured by calipers. Seidell and colleagues (1987) compared the sum of paraumbilical and suprailiac skinfolds with the cross-sectional area of subcutaneous fat measured by computed tomography. They observed high correlations for both men (0.83) and women (0.88).

It is possible that ultrasound measurements may become commonly used in epidemiologic studies, as the method is simple and safe. It should be noted that many of the same limitations of the traditional skinfold technique (sensitivity to the exact placement of the device and the general variation of subcutaneous fat over the body) also applies to the ultrasound method.

Another method to measure thickness of subcutaneous adipose that uses a beam of near-infrared radiation is also commercially available (Jebb and Elia, 1993). However, this approach appears to have limitations similar to traditional skinfold measurements plus relatively low validity.

Validity of Relative Weights, Obesity Indices, and Skinfold Thicknesses as Measures of Body Fat Composition

The validity of epidemiologic measures of body composition can be assessed by comparison with more accurate and precise methods. Because even the best methods are indirect, the choice of an optimal "gold standard," as for dietary intake, is not completely clear. In addition to being highly accurate, it would be desirable that any error associated with the gold standard be independent of error in the method being evaluated so that correlation does not occur simply on the basis of errors that are common to both approaches. Until the present, most studies of validity have employed densitometry (underwater weighing) as the standard method. Although this method is not perfect due, for example, to variation in the bone density of subjects, it is a reasonable choice as any errors should be independent, and it has been in widespread use for decades. It would be reassuring if several of the more sophisticated approaches for measuring obesity (e.g., densitometry, deuterium dilution, electrical conductance, and x-ray absorptiometry) were compared with each other; if very high correlations were observed between them, say on the order of 0.95, this would provide reassurance that they were all providing similarly precise information and could equally serve as standards. A number of studies in which obesity indices and skinfold thicknesses have been compared with densitometry measurements are summarized in Table 10–3. Correlations with obesity indices have ranged from approximately 0.60 to 0.85. Although the correlations with weight/height, weight/height^2, and weight/height^3 are similar, BMI (weight/height^2) tends to be slightly more strongly correlated. The Benn index (weight/height^p) had no higher correlation with densitometry than weight/height^2 in a large study among men (Revicki and Israel, 1986). The correlations of skinfold thicknesses with densitometry have a similar range of coefficients and are not clearly higher than for the obesity indices; it is possible that the use of multiple skinfolds may improve the correlation with densitometry.

Using another approach to evaluate the relative validity of different epidemiologic measures of obesity, Criqui and coworkers (1982) compared various indices with blood triglyceride level, total cholesterol level, blood pressure, and fasting glucose level (which are all known to be related to obesity) among a large population of men and women (Table 10–4). For each outcome, weight/height^2 and relative weight exhibited the strongest correlations, providing evidence that these are the most biologically relevant measures of obesity.

Although the correlations with densitometry seen in Table 10–3 are reasonably high, there are several reasons to believe that they overrepresent the validity of obesity indices and skinfolds in the context of most epidemiologic studies. Because the magnitude of a correlation coefficient is directly related to the degree of variation in the parameter being studied, in this case the between-person variation in obesity, the observed correlation coefficient is applicable only to study populations with a similar variation in obesity. In the published studies, it is often unclear how subjects were selected; however, it seems that they were frequently enriched with an atypically high representation of obese subjects. This would tend to lead to higher correlations than would be observed, given the same degree of accuracy, in a general population. Furthermore, correlations have not been adjusted for age in most published reports. This also overstates the relevant variation in obesity because obesity tends to be strongly correlated with age, and virtually all epidemiologic analyses adjust for age. Not only does BMI increase with age in Western populations, for the same BMI, older persons (and women) tend to have a higher body fat composition (Gallagher et al., 1996). In a study that examined the association of weight-for-height indices with obesity measured by densitometry, Womersley and Durnin (1977) found a correlation of 0.71 for weight/height^2 among men when all ages were combined, but correlations ranging from 0.49 to 0.62 within specific 10-year age groups. Correlations with weight/height^2 were somewhat higher among women, being 0.81 overall, and ranging from 0.64 to 0.91 within specific age groups. Similarly, in a study among 474 men (with unclear basis for selection), Revicki and Israel (1986) found that the correlation between weight/height^2 decreased from 0.71 to 0.58 with adjustment for age. For these reasons, it is difficult to determine the true degree of validity for measures such as BMI in the context of epidemiologic studies on the basis of published data. It seems likely, however, that the correlation with true percent body fat composition in general populations is likely to be on the order of 0.5 or 0.6 for men and perhaps slightly higher for women.

Although use of population samples with atypically large variations in adiposity and the failure to control for basic demographic characteristics may tend to exaggerate correlations between anthropometric indices and percent body fat assessed by densitometry, the validity of these indices may also have been underestimated if percent body fat is not truly the biologically relevant measure of adiposity. Although BMI is typically thought of as an estimate of percent body fat, it is actually more a measure of absolute mass adjusted for height, which is conceptually similar to the residual of

weight adjusted for height described above. Thus, comparisons of BMI with percent body fat do not really compare like with like. Data on percent body fat from densitometry in combination with weight can be used to calculate absolute fat mass, and this can be adjusted for height in regression analysis to create a variable that is conceptually analogous with BMI. Using a large dataset to make such calculations, Spiegelman and colleagues (1992) found remarkably high correlations between BMI and absolute fat mass adjusted for height, even after accounting for age and gender (correlations were between 0.82 and 0.91). Thus, BMI appears to be an excellent measure of fat mass adjusted for height, with somewhat less validity as a measure of percent body fat (the correlations ranged from 0.60 to 0.71).

Using the same dataset, Spiegelman and colleagues (1992) examined the degree to which percent body fat and absolute fat mass adjusted for height, both measured by densitometry, predicted blood pressure and fasting blood glucose level. In both men and women, absolute fat mass adjusted for height appeared to be the better predictor. A variety of other combinations of weight and height as well as measures of body fat distribution were not found to be superior to BMI. These findings suggest that an increase in lean body mass, which would reduce percent body fat, does not offset the adverse effect of excess fat mass. Furthermore, these findings may explain why BMI has been such a strong predictor of health outcomes in a vast epidemiologic literature: It is both a biologically relevant expression of adiposity (apparently better than percent body fat), and it is an excellent measure of adiposity, at least in the young adult and middle-aged population studied by Spiegelman and coworkers.

Despite the excellent performance of BMI as noted above, there are reasons to suspect that this index performs less well in older adults. The basic reason for a high degree of validity is that the vast majority of variation in weight among middle-aged adults of the same gender that is not accounted for by height is fat (muscle builders may be exceptions, but they are rare). However, in the elderly many, but not all, individuals lose substantial amounts of lean body mass (Gallagher et al., 1996), often because of greatly reduced activity. Therefore, variation in lean body mass can contribute to a much greater degree to differences in weight and changes in weight, thus reducing the validity of BMI as a measure of adiposity. Even individuals who maintain the same weight can have substantial changes in adiposity. The observation that men on average do not gain weight appreciably after about age 50, but do greatly expand their abdominal circumference, attests to this redistribution. Indeed, Micozzi and Harris (1990) found that the correlation between BMI and arm fat area tended to decrease with age, and the correlation between BMI and arm muscle area tended to increase with age. Thus, for men over age 65, BMI was similarly correlated with fatness and muscularity, and for women, the correlation with fat area was only modestly greater. It may be that other measures of adiposity will be more appropriate for the elderly. For example, changes in abdominal circumference unequivocally reflect adipose rather than muscle and may thus be a better indicator of overall adiposity than weight or weight indices in some groups.

In summary, on the basis of the previous data, one or two skinfold measurements or any of the obesity indices provide approximately similar estimates of relative body fat composition. Among the obesity indices, weight/height^2 appears at least as good as the others as a measure of relative adiposity and is usually optimal with respect to lack of correlation with height. Although the use of other exponents for height may slightly reduce the correlation with height in a particular population, this rarely outweighs the substantial advantages in comparability among studies that use weight/height^2. Keys and colleagues (1972) have concluded similarly that weight/

Table 10–4. Correlations of height (Ht), weight (Wt), and obesity indices with risk factors in men and women aged 20–79 years, Rancho Bernardo, California, 1972–1974[a]

	Men aged 20–79 yr (n = 2,266)							Women aged 20–79 yr (n = 2,690)						
	Wt	Ht	Wt/Ht	Wt/Ht^2	$\sqrt[3]{Wt/Ht}$	$-Ht/\sqrt[3]{Wt}$	Relative weight	Wt	Ht	Wt/Ht	Wt/Ht^2	$\sqrt[3]{Wt/Ht}$	$-Ht/\sqrt[3]{Wt}$	Relative weight
Age	−0.16	−0.29	−0.09	0.00	−0.14	−0.14	−0.01	0.04	−0.29	0.11	0.19	0.06	0.07	0.17
Cholesterol	0.01	−0.08	0.04	0.07	0.02	0.03	0.07	0.06	−0.17	0.10	0.15	0.07	0.09	0.14
Log triglyceride	0.21	0.01	0.23	0.24	0.22	0.22	0.25	0.19	−0.10	0.23	0.25	0.21	0.20	0.25
Systolic blood pressure	0.01	−0.17	0.07	0.12	0.03	0.02	0.12	0.14	−0.20	0.20	0.25	0.16	0.15	0.25
Diastolic blood pressure	0.14	−0.04	0.17	0.19	0.15	0.15	0.19	0.19	−0.09	0.22	0.24	0.20	0.19	0.25
Fasting plasma glucose	0.08	−0.03	0.10	0.12	0.09	0.09	0.12	0.10	−0.06	0.12	0.13	0.10	0.10	0.13

[a]All correlations of absolute magnitude, 0.07 in women and 0.08 in men, or greater, are significant at $p < 0.001$.

From Criqui et al., 1982.

height^2 is the preferable measure of relative weight in epidemiologic studies. Nevertheless, if height is strongly associated with disease in a particular study, it is important to be certain that the obesity index used is not associated with height. If so, any of a variety of multivariate methods can be used to control for confounding due to height.

The validity of BMI and other obesity indices as measures of relative body composition, represented by the correlations in Table 10–3, is clearly less than perfect. As noted before, part of the reason correlations are not higher is that these represent conceptually different variables. Also, as discussed earlier, only moderate correlation is not the result of error in measuring weight or height; the primary source of error is that these indices reflect the weight of both lean body mass and fat tissue. Bone mass and muscle mass both contribute to the correlation of lean body mass with obesity indices based on weight and height. It is, therefore, important to consider that associations between obesity indices and other variables can be due to differences in lean body mass as well as adiposity. This potential for confounding, however, is due not only to the technical imperfection of the obesity indices as measures of adiposity, but also to the biologic correlation of lean body mass and percent body fat. Thus, the possibility must be considered that an observed association between any measure of obesity and disease is due to an association with lean body mass rather than adiposity.

Dual Energy X-Ray Absorptiometry (DEXA)

In the mid-1980s, DEXA, which uses an x-ray beam with high- and low-energy peaks combined with a whole body scanner, was developed to measure bone mass and has subsequently been used to measure soft tissue composition as well (for further details, see Roubenoff et al., 1993; Lohman, 1996). The method is able to distinguish fat mass, fat-free mass, and bone mineral mass, both for the total body and for specific regions, by the differential absorption of the high- and low-energy x-rays by these tissues. Because the total radiation dose is extremely low, the method can be used for research across all age groups, except pregnant women. It is far easier for participants than underwater weighing. The x-ray and scanning unit is expensive and must be accompanied by software used to estimate body components.

For measurements both within the same day and over months, DEXA provides quite reproducible measurements of body components. For percent body fat, the standard deviation is about 1%, the coefficient of variation (SD/mean × 100%) has been 4% to 7%, depending on the mean body fat in the sample (Lohman, 1996).

The validity of DEXA as assessed by comparison with densitometry appears to be quite high in most populations, but substantial systematic errors have been seen in older and younger individuals (Lohman, 1996). Clark et al. (1993) reported a high correlation (r=0.91) between percent body fat and body density among young white men. In this study, the standard error of the estimate was 3.0% body fat (*s.e.e.* = $\sqrt{\Sigma(Y_o - Y_e)^2/(N - 2)}$, where Y_o is the observed conductivity measurement and Y_e is the conductivity measurement predicted by densitometry; this has been referred to elsewhere as the standard deviation of the residual, see Chapter 6). Hansen et al. (1993) found a similar correlation (0.92) between percent body fat estimated by DEXA and by densitometry among 100 premenopausal women; the standard error of the estimate was 2.4% body fat.

The DEXA method has already found widespread use in clinical studies and may become a standard for other measures of body composition. Whether it or densitometry is a more valid measure of body fat composition is difficult to determine at this point; simultaneous comparisons of both methods with other indicators of body fat would be particularly useful to assess their relative validity.

Bioelectrical Impedance Analysis

In recent years great interest has developed in the use of bioelectrical impedance (or resistance) and conductance measurements to estimate lean body mass and body fat composition (i.e., percent body fat). These measurements are based on the principle that the lean body mass, which consists largely of ions in a water solution, conducts electricity far better than does fat tissue (Van Itallie et al., 1986; Anonymous, 1996; Baumgartner, 1996). Therefore, the resistance (technically impedance in the case of an alternating current) of the body to an electrical current is inversely related to the lean body mass. Such measurements should, therefore, provide the same information as obtained with deuterium oxide or other dilution methods. If the total body mass is known, the fat mass and percent body fat can easily be calculated. Electrical resistance is affected also by body shape, so that correcting measurements for height using empirically derived regression formulas or the ratio height^2/resistance can improve the prediction of body composition. Prediction of lean mass can be further improved by accounting for gender as well as height and weight.*

The bioelectrical impedance method is extremely simple in practice. Electrodes (either two or four) are attached to a person's extremities while recumbent but clothed. A weak radio frequency signal is applied to the electrodes, and the impedance is measured. Usually several measurements are

*A widely accepted prediction equation (Lukaski et al., 1986) is: Fat-free mass (kg) = −4.03 + 0.734 · (Ht^2/R) + 0.116 · (weight) + 0.096 (Xc) + 0.984 · Sex, where *Ht* is in cm, weight is in kg, *Xc* is impedance in ohms, and sex = 0 for F and 1 for M.