

## BMI-Referenced Cut Points for Pedometer-Determined Steps per Day in Adults

C. Tudor-Locke, D.R. Bassett Jr., W.J. Rutherford,  
B.E. Ainsworth, C.B. Chan, K. Croteau, B. Giles-Corti,  
G. Le Masurier, K. Moreau, J. Mrozek, J.-M. Oppert,  
A. Raustorp, S.J. Strath, D. Thompson, M.C. Whitt-Glover,  
B. Wilde, and J.R. Wojcik

**Background:** The goal of this study was to establish preliminary criterion-referenced cut points for adult pedometer-determined physical activity (PA) related to weight status defined by body mass index (BMI). **Methods:** Researchers contributed directly measured BMI and pedometer data that had been collected (1) using a Yamax-manufactured pedometer, (2) for a minimum of 3 days, (3) on ostensibly healthy adults. The contrasting groups method was used to identify age- and gender-specific cut points for steps/d related to BMI cut points for normal weight and overweight/obesity (defined as BMI <25 and  $\geq 25$  kg/m<sup>2</sup>, respectively). **Results:** Data included 3127 individuals age 18 to 94 years (976 men, age =  $46.8 \pm 15.4$  years, BMI =  $27.3 \pm 4.9$ ; 2151 women, age =  $47.4 \pm 14.9$  years, BMI =  $27.6 \pm 6.4$ ; all gender differences NS). Best estimated cut points for normal versus overweight/obesity ranged from 11,000 to 12,000 steps/d for men and 8000 to 12,000 steps/d for women (consistently higher for younger age groups). **Conclusions:** These steps/d cut points can be used to identify individuals at risk, or the proportion of adults achieving or falling short of set cut points can be reported and compared between populations. Cut points can also be used to set

---

Tudor-Locke, Ainsworth, Mrozek, and Wilde are with the Dept of Exercise and Wellness, Arizona State University, Mesa, AZ 85212. Bassett and Thompson are with the Dept of Exercise, Sport, and Leisure Studies, The University of Tennessee, Knoxville, TN 37996. Rutherford is with the Dept of Exercise and Sport Science, Eastern Kentucky University, Richmond, KY 40475. Chan is with the Dept of Physiology, University of Alberta, Edmonton, AB T6G 2R3 Canada. Croteau is with the Dept of Exercise, Health, and Sport Sciences, University of Southern Maine, Gorham, ME 04038. Giles-Corti is with the School of Population Health, The University of Western Australia, Crawley, Australia 6009. Le Masurier is with Malapina University-College, Nanaimo, BC, Canada V9R 5S5. Moreau is with the Division of Geriatric Medicine, University of Colorado at Denver and Health Sciences Center, Aurora, CO 80045. Oppert is with the Dept of Nutrition, Pitie-Salpetriere Hospital, 75013 Paris, France. Raustorp is with the School of Human Science, University of Kalmar, SE 391 82 Kalmar, Sweden. Strath is with the Dept of Human Movement Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201. Whitt-Glover is with the Dept of Epidemiology and Prevention, Wake Forest University School of Medicine, Winston-Salem, NC 17157. Wojcik is with the Dept of Health and Physical Education, Winthrop University, Rock Hill, SC 29733.

intervention goals, and they can be referred to when evaluating program impact, as well as environmental and policy changes.

**Keywords:** walking, obesity, exercise, objective monitoring

Pedometers are simple and inexpensive body-worn motion sensors that are increasingly being used by researchers and practitioners to measure and motivate physical activity (PA) behaviors. They have also become widely available to the general public and are being distributed through such everyday channels as cereal boxes and restaurant meals. To interpret pedometer-determined PA outputs, however, researchers and practitioners, as well as the lay public, require steps/d indices associated with important health-related outcomes (eg, obesity, hypertension, etc) and/or healthy levels of PA (ie, steps/d equivalents of accepted public health recommendations).<sup>1</sup> Measurement tools without interpretable thresholds, benchmarks, or cut points are of limited use.

A traditional approach to setting behavior or performance guidelines is the norm-referenced approach, whereby individual differences are compared to normative or expected values.<sup>2</sup> An early attempt to assemble expected values for steps/d was based on aggregated data from a number of studies (diverse in their pedometer-brand use) that concluded that healthy adults take between 7000 and 13,000 steps/d.<sup>3</sup> The oft-cited goal of 10,000 steps/d, considered indicative of an active lifestyle and associated with health benefits,<sup>4,5</sup> lies within this range. In addition, 10,000 steps/d has been generally accepted by the media and commercial entities without any authoritative endorsement to date.

Another approach has been to convert existing public health recommendations to a steps/d metric. The benefit of this approach is that steps/d guidelines can serve to translate the recommendations into a form that is easily understood by the general public. Widely accepted, evidence-based adult public health guidelines promote  $\geq 30$  minutes of at least moderate-intensity daily PA.<sup>6</sup> Consistent evidence supports that 30 minutes of at least moderate-intensity walking is equivalent to 3000 to 4000 steps in adults.<sup>7,8</sup> Furthermore, there is preliminary support for a stepping rate of 100 steps/min to represent the lower bound of moderate-intensity walking in adults.<sup>9</sup> To be considered a true translation of public health guidelines, however, these 3000 to 4000 steps should be of at least moderate intensity (ie,  $\geq 100$  steps/min), accumulated in at least 10-minute bouts, and be taken *over and above* some threshold level steps/d indicative of sedentarism. A value of  $\leq 5000$  steps/d has been proposed as a sedentary lifestyle index.<sup>1</sup> Summing this value and the supplemental steps/d minimally representative of public health recommendations produces a putative floor value of 8000 to 9000 steps/d (again, applied with the caveats listed above). Although acknowledging health benefits associated with  $\geq 30$  minutes of moderate-intensity daily activity, the Institute of Medicine issued a 2002 report<sup>9</sup> that recommended twice this duration (ie, at least 60 minutes of moderate-intensity daily activity) to prevent weight gain. Following the same additive process just described (ie, a sedentary level of 5000 steps/d plus the step equivalent of 60 minutes of moderate-intensity activity, 6000–8000 steps) produces a total value of 11,000 to 13,000 steps/d for weight maintenance.

More appropriately, cut points for pedometer-determined PA should emerge from data related to important health outcomes (eg, a healthy body weight category

as indicated broadly by a BMI < 25 kg/m<sup>2</sup>). This general strategy for setting guidelines is known as the criterion-referenced approach,<sup>1</sup> and it considers a specific health outcome or a disease risk factor (ie, the criterion) to be the critical factor in setting cut points. This approach has been previously used to set BMI-referenced steps/d cut points in children<sup>3</sup> but not yet in adults. The purpose of the current study, therefore, was to describe the collaborative process involved in establishing an international pedometer amalgamated data set and to establish preliminary criterion-referenced cut points for PA using pedometer-assessed steps/d related to normal-weight (BMI < 25 kg/m<sup>2</sup>) versus overweight/obese (BMI ≥ 25 kg/m<sup>2</sup>) adults as determined by directly measured height and weight.

## Methods

### Procedures

During 2004, international pedometer researchers contributed data that had been collected (1) using a specific pedometer brand (see rationale below); (2) for a minimum of 3 days (see rationale below); (3) on ostensibly healthy adults who were not recruited specifically on PA, body-weight category, or disease status (eg, diabetes) variables. In addition, direct and complete measures of height and weight were required to compute BMI. An initial attempt to collect race/ethnicity was abandoned when it was realized that these variables were inconsistently collected (ie, definitions and groupings varied widely), if at all. For example, it is illegal to collect such data in France. The first author obtained Institutional Review Board approval through Arizona State University for this secondary data analysis. Local institutional review board approval was also secured by each of the individual collaborators.

### Instrument

Only those data collected using the Yamax pedometer (Yamax Corporation, Tokyo, Japan) were considered for inclusion. This pedometer has been the instrument of choice for most scientific studies to date and has performed well in brand-to-brand comparisons.

### Monitoring Frame

Three days has been recommended as a minimum monitoring frame to ascertain an estimate of weekly steps/d in adults<sup>3</sup>; the intraclass correlation for any 3-day combination relative to the week-based mean in that study ranged from .86 to .91. Records of pedometer time worn during the day are unavailable for the independent samples forming this amalgamated data set; a minimal timeframe for daily wear has not been specified for pedometry.

### Data Management

The first author collected and managed all donated data and ensured that variables were consistently represented. All incoming data, regardless of the original database

software used, were translated to SPSS V 12 using DBMS/COPY software. To prepare for this process, potential collaborators were asked to submit a Description of Data form before any data transfer. The form included questions about pedometer used (ie, brand, model, manufacturer/distributor, city, etc), number and gender of subjects, age range, number of days of monitoring, and database used. Researchers were asked to supply a data dictionary, that is, a complete accounting of all relevant variable names and their definitions.

### Statistical Analyses

Scatter plots and Pearson correlations were used to describe sex-specific bivariate associations between steps/d and (1) age and (2) BMI. The contrasting-groups method<sup>6</sup> was used to establish criterion-referenced cut points to identify adult age- and gender-specific cut points for steps/d related to accepted BMI cut points for normal weight and overweight/obese. A large and diverse sample was required because the contrasting-groups method recommends at least 100 individuals in each of the 2 designated BMI categories for every subgroup studied (eg, age and gender subgroups). Stratifying by both age and gender is justified by the well-known influence of these factors on steps/d.<sup>7</sup> Specific age categories were dictated by the size of the gender-specific samples and the 100-individual minimum (this was not possible for older men only). Otherwise, the gender- and age-specific analysis samples included all subjects identified as overweight/obese and a randomly drawn number-matched selection of gender- and age-matched normal-weight subjects. A further division and analysis of the data into overweight versus obese was not possible with this data set because of insufficient numbers. Representativeness of the amalgamated analysis sample was verified against the source data.

A distribution of average steps/d for each gender and age group was computed and smoothed to 1000 step/d increments as was previously performed with a similar analysis of children.<sup>3</sup> The contrasting-groups method of establishing cut points is predicated on the presence of dichotomous groups (ie, defined herein as normal-weight and overweight/obese adults). The first task in the method is to plot frequency distributions of steps/d for both groups within each gender- and age-specific stratum. The intersection of the frequency distributions is used as the hypothetical cut point for steps/d. Considerable overlap is often observed between the 2 groups, indicating that there are "false normal weight" adults with high BMI who exceeded the steps/d cut point (*false masters* is the original terminology presented in the contrasting-groups method or *false positives* in the epidemiological literature; it is modified slightly herein to reflect more appropriate use of language in terms of body-weight category) and "false overweight/obese" adults with low BMI who did not exceed the cut point (*false non-masters* is the original terminology presented in the contrasting-groups method or *false negatives* in the epidemiological literature). The overlapping of scores is common in studies of this kind and represents a limitation to setting criterion-referenced cut points using this method. Regardless, scores near the hypothetical cut point can be analyzed by interpreting several indices as described later and elsewhere<sup>3</sup> generally and originally and in much more detail (including specific calculation strategies too detailed to include here) by Berk.<sup>8</sup>

The index probability of correct decisions represents the extent to which the cut point maximizes the probability of correct classifications of true normal weight and true overweight/obese versus the probability of incorrect classifications of false normal weight and false overweight/obese. This index also represents the extent to which the probability of normal-weight adults who exceed the cut point and overweight/obese adults who do not exceed the cut point is minimized. A high probability index is, therefore, desirable.

Misclassification errors estimate the likelihood of incorrectly classifying someone's BMI status (ie, normal weight versus overweight/obese) on the basis of steps/d. Risks of misclassification can be weighted to account for the seriousness of each type of error. In this study, the risks of each type of error were considered to be equal and were, therefore, given equal weight.

The phi (validity) coefficient describes the relationship between BMI and steps/d. Phi coefficients are interpreted as other correlation coefficients, with higher values indicative of higher validity.

Utility analysis is an estimate of the usefulness of the cut point as measured by the gains and losses associated with classification and misclassification probabilities. Specifically, expected utility is the gain associated with correct classification and is equal to the sum of the products of the true probabilities and their respective utility values. Expected disutility is the loss associated with misclassifying a subject on the basis of BMI status. This loss is a function of the seriousness of type I and type II errors. Expected disutility is calculated by summing the proportions of both false normal weight and false overweight/obese, after first assigning whatever weightings are appropriate to the misclassifications. In this study, each misclassification was assigned equal weight. Expected maximal utility is a composite measure of test usefulness and is determined by summing the expected utility and disutility and multiplying the sum by the sample size of the combined criterion groups.

Selection of the final cut point is based on a considered reflection of all statistical indices, or a heavier weighting might be assigned to 1 or more factors. A selected cut point should ideally exhibit a high probability of correct decisions, low probabilities of type I and type II errors, a high phi coefficient, high expected utility, low expected disutility, and high expected maximal utility. This specific constellation of indices seldom occurs in practice, and the ultimate cut point is, therefore, a best estimate with acknowledged limitations.

## Results

Ten researcher groups from the USA contributed data. Four others contributed data from Australia, Canada, France, and Sweden, respectively, which together represent 40% of the sample. Previously published papers containing these data are compiled in the appendix. The final amalgamated data set represented 3127 individuals age 18 to 94 years (976 men, age =  $46.8 \pm 15.4$  years, BMI =  $27.3 \pm 4.9$ ; 2151 women, age =  $47.4 \pm 14.9$  years, BMI =  $27.6 \pm 6.4$ ; all gender differences NS). Men took  $8548 \pm 3699$  steps/d, and women took  $7494 \pm 3459$  steps/d ( $t = 5.527$ ,  $P < .0001$ ). Steps/d was inversely correlated with age and BMI in men ( $r = -.33$  and  $r = -.28$ , respectively) and women ( $r = -.30$  and  $r = -.31$ , respectively), all  $P < .01$ .

The gender-specific age groups that emerged from the data (applying the analysis requirements described previously) were 18 to 50.9 and 51 to 88 years of age for men and 18 to 39.9, 40 to 49.9, 50 to 59.9, and 60 to 94 years of age for women. The greater number of age groups for women versus men was a result of their relatively larger numbers in the amalgamated data set. Mean  $\pm$  SD steps/d for the whole sample are presented in Table 1 for each gender and age group, stratified for BMI status. Differences in steps/d were significant between normal-weight and overweight/obesity categories for each gender-age stratum, satisfying a fundamental requirement for the application of the contrasting-groups method.<sup>5</sup>

A summary of the analytical process (ie, probability of correct decisions, misclassification of errors, validity coefficient, utility analysis) undertaken to evaluate and select step/d cut points (indicated by boldface font in the tables) for each age group is presented in Table 2 (men) and Table 3 (women). Based on best-estimated criterion-referenced cut points indicated for each gender and age group, the cut point for steps/d associated with normal weight versus overweight/obesity (defined by BMI cut points) ranged from 11,000 to 12,000 for men and 8000 to 12,000 for women (consistently higher for younger age groups).

To illustrate the application of the contrasting-groups method graphically, a 2x2 classification of women age 50 to 59.9 years is presented in Figure 1. Distribution of steps/d of normal-weight and overweight/obese strata are shown along the abscissa, and the frequency of the scores is shown along the ordinate. The distributions of the 2 groups are the primary determinants of the extent to which steps/d can accurately classify individuals as true normal weight and true overweight/obese. The degree of accuracy is largely a function of the amount of overlap between these distributions. Complete overlap suggests that the index (steps/d) is useless for classification purposes. No overlap would suggest the index could classify individuals with perfect accuracy. Naturally, some overlap is to be expected, and this is exemplified in Figure 1 by the positive skewness of both distributions.

The point at which the distributions intersect is the best cut point of those considered. In Figure 1, this score is approximately 6700 steps/d. Scores above this cut point increase type I misclassification error (false overweight) and decrease

**Table 1 Mean  $\pm$  SD Steps per Day for Men and Women Stratified by Weight Status**

Age (y)	Normal weight, steps/d (n)	Overweight/Obese, steps/d (n)	t	P
<b>Men</b>				
18-50.9	10,084 $\pm$ 3503 (217)	8802 $\pm$ 3709 (367)	4.117	.000
51-88	8938 $\pm$ 3451 (83)	6924 $\pm$ 3318 (296)	4.843	.000
<b>Women</b>				
18-39.9	9399 $\pm$ 3348 (320)	7562 $\pm$ 3453 (321)	6.839	.000
40-49.9	8726 $\pm$ 3124 (233)	7177 $\pm$ 3000 (320)	5.891	.000
50-59.9	8289 $\pm$ 3670 (143)	6580 $\pm$ 3324 (357)	5.042	.000
60-94	7021 $\pm$ 3362 (156)	5362 $\pm$ 2801 (267)	5.196	.000

Table 2a Summary of the Analytical Process to Evaluate Steps/d Cut Points in Men Age 18 to 50.9 y (n = 226)

Steps/d	Probability of correct decisions	Misclassification error: type I, type II (false overweight/obese, false normal weight)	Validity coefficient (phi)	Utility analysis		
				Expected utility	Expected disutility	Expected maximum utility
9000	.6	.26, .24	-.05	.5	-.74	-.0005
10,000	.47	.32, .18	.01	.5	-.68	-.0004
11,000	.37	.35, .15	.04	.5	-.65	-.0003
<b>12,000*</b>	.27	<b>.39, .11</b>	.05	.5	-.61	-.0002
13,000	.21	.42, .08	.055	.5	-.58	-.0002

\*Boldface font indicates best cut point.

Table 2b Summary of the Analytical Process to Evaluate Steps/d Cut Points in Men Age 51 to 88 y (n = 220)

Steps/d	Probability of correct decisions	Misclassification error: type I, type II (false overweight/obese, false normal weight)	Validity coefficient (phi)	Utility analysis		
				Expected utility	Expected disutility	Expected maximum utility
8000	.44	.36, .14	.04	.5	-.64	-.0006
9000	.31	.41, .09	.12	.5	-.59	-.0004
10,000	.21	.45, .09	.16	.5	-.55	-.0002
<b>11,000*</b>	.15	<b>.48, .02</b>	.21	.5	-.52	-.0001
12,000	.11	.48, .02	.19	.5	-.52	-.00008
13,000	.09	.48, .009	.2	.5	-.5	-.00004
14,000	.06	.5, .005	.19	.5	-.5	-.00002

\*Boldface font indicates best cut point.

**Table 3a Summary of the Analytical Process to Evaluate Steps/d Cut Points in Women Age 18 to 39.9 y (n = 321)**

Steps/d	Probability of correct decisions	Misclassification error: type I, type II (false overweight/obese, false normal weight)	Validity coefficient (phi)	Utility analysis		
				Expected utility	Expected disutility	Expected maximum utility
8000	.53	.30, .20	-.01	.5	-.7	-.0003
9000	.41	.35, .15	.04	.5	-.65	-.0002
10,000	.31	.40, .10	.09	.5	-.6	-.0002
11,000	.22	.42, .08	.09	.5	-.58	-.0001
<b>12,000*</b>	.17	.45, .05	.13	.5	-.55	-.00007
13,000	.12	.47, .03	.13	.5	-.53	-.00004
14,000	.07	.48, .02	.09	.5	-.52	-.00003
15,000	.53	.30, .20	-.01	.5	-.7	-.0003

\* Boldface font indicates best cut point.

**Table 3b Summary of the Analytical Process to Evaluate Steps/d Cut Points in Women Age 40 to 49.9 y (n = 298)**

Steps/d	Probability of correct decisions	Misclassification error: type I, type II (false overweight/obese, false normal weight)	Validity coefficient (phi)	Utility analysis		
				Expected utility	Expected disutility	Expected maximum utility
8000	.44	.39, .11	.06	.5	-.61	-.0004
9000	.37	.42, .08	.13	.5	-.58	-.0003
10,000	.29	.45, .05	.19	.5	-.55	-.0002
<b>11,000*</b>	.18	.47, .03	.21	.5	-.53	-.0001
12,000	.13	.48, .02	.19	.5	-.52	-.00008
13,000	.1	.48, .02	.17	.5	-.52	-.00007

\* Boldface font indicates best cut point.

Table 3c Summary of the Analytical Process to Evaluate Steps/d Cut Points in Women Age 50 to 59.9 y (n = 366)

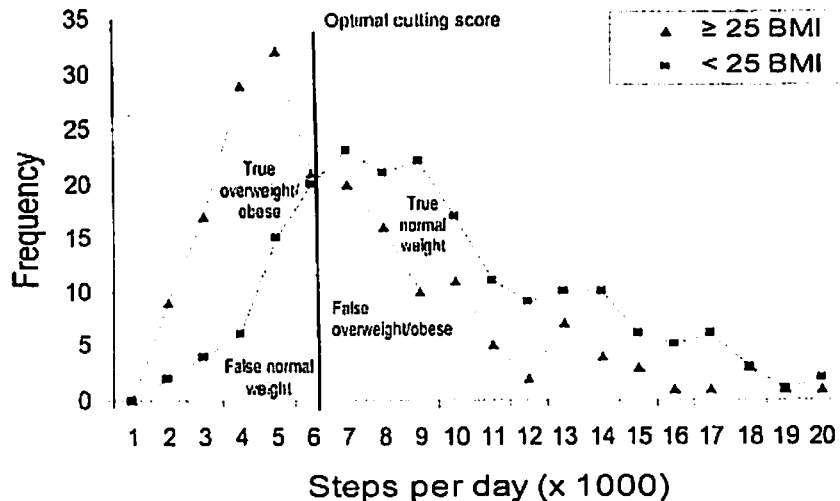
Steps/d	Probability of correct decisions	Misclassification error: type I, type II (false overweight/obese, false normal weight)	Validity coefficient (phi)	Utility analysis		
				Expected utility	Expected disutility	Expected maximum utility
7000	.45	.36, .14	.04	.5	-.64	-.0004
8000	.4	.40, .10	.09	.5	-.6	-.0003
9000	.28	.43, .07	.14	.5	-.57	-.0002
<b>10,000*</b>	.2	.46, .04	.18	.5	-.54	-.0001
11,000	.17	.46, .04	.18	.5	-.54	-.0001
12,000	.14	.47, .03	.15	.5	-.53	-.00009

\* Boldface font indicates best cut point.

Table 3d Summary of the Analytical Process to Evaluate Steps/d Cut Points in Women Age 60 to 94 y (n = 214)

Steps/d	Probability of correct decisions	Misclassification error: type I, type II (false overweight/obese, false normal weight)	Validity coefficient (phi)	Utility analysis		
				Expected utility	Expected disutility	Expected maximum utility
6000	.39	.36, .14	.06	.5	-.64	-.0006
7000	.25	.42, .08	.1	.5	-.58	-.0004
<b>8000*</b>	.18	.44, .06	.11	.5	-.56	-.0003
9000	.14	.45, .05	.09	.5	-.55	-.0002
10,000	.09	.47, .03	.07	.5	-.53	-.0001

\* Boldface font indicates best cut point.



**Figure 1** — Example of contrasting-groups classification of steps/d cut points based on normal weight versus overweight/obese (BMI < 25 versus  $\geq 25$ , respectively) for women age 50 to 59.9 years.

type II misclassification error (false normal weight). As stated previously, the positive skewness and obvious overlap is suggestive of considerable error with regard to classification decisions. Identification of the best cut point, therefore, requires the contrasting-groups analysis be performed, the results of which are presented in Tables 2 (men) and 3 (women).

By way of example, as indicated in Table 3, the best cut point for women ages 50 to 59.9 years is 10,000 steps/d. This score was selected because its validity, expected utility, expected disutility, and expected maximum utility indices were best when compared with other cut points. It should be emphasized, however, that these values are relatively low and suggest the potential for considerable error in classification decisions.

Interpretation of the contrasting-groups analysis might be made as follows. The cut point that best predicts group (ie, normal weight versus overweight/obese) classification for women age 50 to 59.9 years is 10,000 steps/d. The probability of correctly classifying women age 50 to 59.9 years as normal weight versus overweight/obese using this cut point is only .20 or 20%. There is a 46% chance of misclassifying women age 50 to 59.9 years as false overweight/obese but only a 4% chance of misclassifying them as false normal weight. The validity coefficient of .18 suggests that the cut point is only 18% accurate in estimating the correct classification. Utility analysis is a more thorough analysis of the costs or losses associated with misclassification of individuals. In general, higher values are indicative of lower cost and are, therefore, desirable. For a more complete discussion of the utility-analysis concept, see the seminal paper on the contrasting-groups method.<sup>8</sup>

## Discussion

Overall, the average steps/d taken by this combined adult sample is within the range of expected values and is consistent with known age- and gender-specific differences.<sup>9</sup> Furthermore, the correlations observed between steps/d and both age and BMI are within published values.<sup>7</sup> These findings support the contention that the amalgamated data set is representative of the general population, a crucial consideration to setting criterion-referenced cut points that are externally valid. Despite this, individual samples submitted cannot be interpreted as representative of their source countries, and there was, therefore, no attempt to present the data in this manner. A diverse, heterogeneous sample that depicts a range of behavior and weight status is actually a requirement of the contrasting-group method for setting cut points, and, therefore, the international characteristic of this sample is an asset, not a limitation.

Women outnumbered the men in this data set by more than twofold. The reduced number of men restricted our ability to examine finer age gradations. Conclusions about the associated cut points are consequently limited. Although additional numbers of both genders (and all ages) are required to increase assurance in this data set, efforts to address the gender imbalance by selectively recruiting men in future collection attempts will ultimately build a stronger data set and allow for more confident exploration of age-specific cut points within the male subgroup. Another gender-related limitation of the current study is that because we could not control for the use of hormone therapy in postmenopausal women, we cannot rule out a potential confounding relation between body weight and steps/d in women. This was an amalgamation of different data sets, and we were limited to analyzing only those variables commonly collected. Future studies might need to examine the relation between hormone therapy use, steps/d, and weight gain in women.

Although criterion-based approaches to setting cut points are considered improvements over norm-based approaches in terms of relating behaviors or performances to health-related outcomes,<sup>2</sup> misclassification error is an inevitable limitation. As described above, the 2 possible misclassification errors that can occur are false normal weight and false overweight/obese. There will be inexorable exceptions when applying any set cut point because BMI is obviously a product of numerous interacting factors in addition to PA (again, as evidenced by its low correlation herein with steps/d). A number of the computed indices were not in expected directions or magnitudes, and the final selection process was tempered with vigilance.

Applying due caution, however, the best criterion-referenced cut-point estimates ranged from 11,000 to 12,000 steps/d for men and 8,000 to 12,000 steps/d for women. A pattern of reducing cut points with age was most apparent with women (because there were more age categories possible), but a consistent trend was observed with men. The reduced need for steps/d to maintain a normal-weight BMI level with age is likely influenced by age-related declines in energy intake.<sup>10</sup> For men and women in the youngest age groups, 12,000 steps/d was associated with a BMI in the normal category. This value is not altogether different from that computed previously by converting the Institute of Medicine's<sup>11</sup> recommendation for weight maintenance or prevention of weight gain (directly relevant to

the select health-related criterion used herein). It is also within compiled norm-referenced values.<sup>9</sup> It is not consistent, however, with ACSM-CDC public health recommendations<sup>12</sup> based originally on all-cause mortality as an outcome (that is, not focused on weight maintenance or prevention of weight gain). Furthermore, 12,000 steps/d appears to be in direct disagreement with other authors' speculations that an increase of 2000 to 2500 steps/d (approximately 100/kcals expended daily) might be sufficient to stave off weight gain.<sup>13</sup> Although it appears to be a potentially palatable message, these results suggest that it will be insufficient to achieve the desired public health outcome. Moreover, we note that 10,000 steps/d was only indicated as a BMI-referenced cut point for 50- to 59.9-year-old women in this data set.

In 2004, Tudor-Locke and Bassett Jr<sup>1</sup> proposed preliminary pedometer-determined PA cut points for healthy adults based on the published literature: (1) <5000 steps/d (sedentary), (2) 5000 to 7499 steps/d (low active), (3) 7500 to 9999 (somewhat active), (4)  $\geq 10,000$  to 12,499 (active), and  $\geq 12,500$  steps/d (highly active). The zone delimited by 7500 to 10,000 steps/d (described as somewhat active) is amassing support as evidence continues to accrue that health benefits can be realized and that accepted public health recommendations are achievable within this zone.<sup>14-16</sup> It is also congruent with the strict conversion of these recommendations to steps/d values (described previously). Known dose-response relationships dictate that additional benefits are realized at higher volumes of PA, and the likelihood of participating in activity of recommended intensity levels is increased at higher levels of steps/d.<sup>17</sup> With the stated limitations of setting criterion-referenced cut points, we found that 12,000 steps/d appears to be associated specifically with a healthy body-weight category (as broadly indicated by a BMI of  $<25$  kg/m<sup>2</sup>) in younger women and men. Although a BMI-referenced steps/d cut point is useful for setting an anchor within the zone hierarchy of steps/d (outlined previously), from an individual intervention perspective, it is best interpreted as assistive rather than prescriptive.

In this initial analysis of these comparable merged data, we used accepted analytical procedures to establish BMI-referenced cut points for pedometer-determined PA in adults. Steps/d cut points can be useful for guiding screening, surveillance, intervention, and evaluation. For example, these steps/d cut points can be used to identify individuals at risk, or the proportion of adults achieving or falling short of set cut points can be reported and compared between populations. Cut points can also be used to set intervention goals, and they can be referred to when evaluating program impact, as well as environmental and policy changes. Despite their apparent usefulness, however, it is important to emphasize that the specific cut points established herein are preliminary and require rigorous cross-validation (this was not an option herein because the number of subjects per cell was too small) and additional evaluation before any form of universal acceptance. Although we took care to set inclusion criteria that ensure comparability of the distinct data samples that contributed to the final amalgamated data set, we make no assertion that the data are otherwise accurate or even as representative of what could be obtained from a random sampling. It is important also to state that these data are cross-sectional, and conclusions about causality cannot be inferred. Longitudinal study designs (both observational and intervention) are required to more clearly delineate a causal dose-response relationship between steps/d and body-weight categories.

### Acknowledgments

The Western Australian Premier's Physical Activity Task Force is acknowledged for making available its pedometer data. Dr Giles-Corti was Chair of its Evaluation and Monitoring Working Group that oversaw the collection of these data. The contribution of Mr Gavin McCormack is gratefully acknowledged. The data contributed by Dr Barbara Ainsworth were collected from research studies completed at the University of South Carolina Prevention Research Center and Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC.

### References

1. Tudor-Locke C, Bassett DR Jr. How many steps/day are enough? preliminary pedometer indices for public health. *Sports Med.* 2004;34(1):1-8.
2. Safrit MJ, Wood TM. *Introduction to Measurement in Physical Education and Exercise Science.* 3rd ed. St Louis, MO: Times Mirror/Mosby; 1995.
3. Tudor-Locke C, Pangrazi RP, Corbin CB, et al. BMI-referenced standards for recommended pedometer-determined steps/day in children. *Prev Med.* 2004;38(6):857-864.
4. Crouter SC, Schneider PL, Karabulut M, Bassett DR Jr. Validity of 10 electronic pedometers for measuring steps, distance, and energy cost. *Med Sci Sports Exerc.* 2003;35(8):1455-1460.
5. Tudor-Locke C, Burkett L, Reis JP, Ainsworth BE, Maccra CA, Wilson DK. How many days of pedometer monitoring predict weekly physical activity in adults? *Prev Med.* 2005;40(3):293-298.
6. Safrit MJ. *Introduction to Measurement in Physical Education and Exercise Science.* St Louis, MO: Mosby College Publishing; 1986.
7. Tudor-Locke C, Williams JE, Reis JP, Pluto D. Utility of pedometers for assessing physical activity: construct validity. *Sports Med* 2004;34(5):281-91.
8. Berk RA. Determination of optimal cutting scores in criterion-referenced measurement. *Journal of Experimental Education* 1976;45(4-9).
9. Tudor-Locke C, Myers AM. Methodological considerations for researchers and practitioners using pedometers to measure physical (ambulatory) activity. *Res Q Exerc Sport.* 2001;72(1):1-12.
10. Wakimoto P, Block G. Dietary intake, dietary patterns, and changes with age: an epidemiological perspective. *J Gerontol A Biol Sci Med Sci.* 2001;56(Spec No 2):65-80.
11. Institute of Medicine. *Dietary Reference Intakes for Energy, Carbohydrates, Fiber, Fat, Protein and Amino Acids (Macronutrients).* Washington, DC: National Academy of Sciences; 2002.
12. Pate RR, Pratt M, Blair SN, et al. Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *J Am Med Assoc.* 1995;273(5):402-407.
13. Hill JO, Wyatt HR, Reed GW, Peters JC. Obesity and the environment: where do we go from here? *Science.* 2003;299(5608):853-855.
14. Sugiura H, Kajima K, Mirbod SM, Iwata H, Matsuoka T. Effects of long-term moderate exercise and increase in number of daily steps on serum lipids in women: randomised controlled trial [ISRCTN21921919]. *BMC Womens Health* 2002;2(1):3.
15. Tudor-Locke C, Bell RC, Myers AM, et al. Controlled outcome evaluation of the First Step Program: a daily physical activity intervention for individuals with type II diabetes. *Int J Obes Relat Metab Disord.* 2004;28(1):113-119.

16. Jordan AN, Jurca GM, Tudor-Locke C, Church TS, Blair SN. Pedometer indices for weekly physical activity recommendations in postmenopausal women. *Med Sci Sports Exerc.* 2005;37(9):1627-1632.
17. Le Masurier GC, Sidman CL, Corbin CB. Accumulating 10,000 steps: does this meet current physical activity guidelines? *Res Q Exerc Sport.* 2003;74(4):389-394.

## Appendix

### Previously Published Papers Based on Portions of the Assembled International Amalgamated Data Set

- Anderson ES, Wojcik JR, Winett RA, Williams DA. Social cognitive determinants of physical activity: the influence of social support, self-efficacy, outcome expectations, and self-regulation. *Health Psych.* 2006;25(4):510-520.
- Chan CB, Ryan DAJ, Tudor-Locke C. Health benefits of a pedometer-based physical activity intervention in sedentary workers. *Prev Med.* 2004;39(6):1215-1222.
- Chan CB, Spangler E, Valcourt J, Tudor-Locke C. Cross-sectional relationship of pedometer-determined ambulatory activity to indicators of health. *Obes Res.* 2003;11(12):1563-1570.
- Croteau KA. A preliminary study on the impact of a pedometer-based intervention on daily steps. *Am J Health Promot.* 2004;18:217-220.
- Kettaneh A, Oppert JM, Heude B, et al. Changes in physical activity explain paradoxical relationship between baseline physical activity and adiposity changes in adolescent girls: the FLVS II study. *Int J Obes.* 2005;29:586-593.
- Moreau KL, DeGarmo R, Langley J, et al. Increasing daily walking lowers blood pressure in postmenopausal women with borderline to stage 1 hypertension. *Med Sci Sports Exerc.* 2001;33(11):1825-1831.
- Raustorp A, Mattsson E, Svensson K, Ståhle A. Physical activity, body composition, and physical self-esteem: a three year follow-up study among adolescents in Sweden. *Scan J Med Sci Sports.* 2006;16:258-266.
- Thompson DL, Rakow J, Perdue SM. Relationship between accumulated walking and body composition in middle-aged women. *Med Sci Sports Exerc.* 2004;36(5):911-914.
- Whitt M, Kumanyika S, Bellamy S. Amount and bouts of physical activity in a sample of African-American women. *Med Sci Sports Exerc.* 2003;35(11):1887-1893.
- Winett RA, Anderson ES, Wojcik JR, Winett SG, Bowden T. Guide to health: nutrition and physical activity outcomes of a group-randomized trial of an Internet-based intervention in churches. *Ann Behav Med.* 2007;33(3):251-261.