ies: modification of the Astrand-Rhyming test. Med Sci Sports Exer 14: 335–338 (1982).
19. Cooper, K. H.: The aerobics way. M. Evans and Company, Inc., New York, 1978.
20. Behnke, A. R., and Wilmore, J. H.: Evaluation and regulation of body build and composition. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1972.
21. Keys, A., et al.: Indices of relative weight and obesity. J Chronic Dis 25: 329–343 (1974).
22. Farquhar, J. W.: The American way of life need not be hazardous to your health. W. W. Norton and Company, New York, 1978.
23. Clarke, H. H.: Improvements of objective strength tests of muscle groups by cable tension methods. Res Q 21: 399 (1950).
24. McArdle, W. D., Katch, F. I., and Katch, V. L.: Exercise physiology: energy, nutrition and human performance. Lea and Febiger, Philadelphia, 1981.
25. Larson, L. A., editor: Fitness, health, and work capacity: international standards for assessment. MacMillan Publishing Co., Inc., New York, 1974.
26. Blair, S. N., Falls, H. B., and Pate, R. R.: A new physical fitness test. Physician Sports Med 11: 87–95 (1983).
27. Leighton, J. R.: Instrument and technic for measurement of range of joint motion. Arch Phys Med Rehabil 38: 24–28 (1955).

Assessment of Physical Activity in Epidemiologic Research: Problems and Prospects

RONALD E. LAPORTE, PhD
HENRY J. MONTOYE, PhD
CARL J. CASPERSEN, PhD, MPH

Dr. LaPorte is an Associate Professor in the Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, 130 DeSoto St., Pittsburgh, PA 15261. Dr. Montoye is a Professor in the Department of Physical Education and Dance, School of Education, University of Wisconsin-Madison, Madison, WI 53706. Dr. Caspersen is an Epidemiologist in the Behavioral Epidemiology and Evaluation Branch, Division of Health Education, Center for Health Promotion and Education, Centers for Disease Control, Atlanta, GA 30333. Richard A. Washburn, PhD, Research Fellow, Department of Epidemiology, School of Public Health, University of Pittsburgh, made helpful suggestions.

Tearsheet requests to Dr. LaPorte.

Synopsis

More than 30 different methods have been used to assess physical activity. These methods can be grouped into seven major categories: calorimetry, job classification, survey procedures, physiological markers, behavioral observation, mechanical and electronic monitors, and dietary measures. No single instrument fulfills the criteria of being valid, reliable, and practical while not affecting behavior. The instruments that are very precise tend to be impractical on a population basis. Surveys are the most practical approach in large-scale studies, although little is known about their reliability and validity. Studies employing objective monitoring through heart rate, movement sensors, and doubly labeled water procedures appear promising, but are still experimental and costly. Despite the difficulty of measurement, relatively strong association has been found between physical activity and health, suggesting that, with improvements in assessment techniques, even stronger associations should be seen.

A CRITICAL FACTOR FOR EPIDEMIOLOGIC RESEARCH is the accurate assessment of the variables under study. For the cardiovascular risk factors of smoking, serum cholesterol, and blood pressure, there are standardized techniques for assessment of the factors that provide consistency of measurement and definition across studies. However, with research on physical activity this has not been the case.

This review considers seven major categories of physical activity assessment procedures that have been used in various settings (table 1) and evaluates their potential for use in epidemiologic studies with respect to four important criteria:

- To be valid, the instrument must measure what it is intended to measure.
- To be reliable, the instrument must consistently give the same results under the same circumstances. If the instrument is reliable and valid, it is also accurate.
- To be practical, the instrument must have acceptable costs to both the investigator and the participant.

Table 1. Characteristics of physical activity assessment procedures

Activity assessment procedure	Group: Size	Group: Age	Study costs[1]: Money	Study costs[1]: Time	Subject costs[1]: Time	Subject costs[1]: Effort	Probability of interfering[1]	Acceptability: Personal	Acceptability: Social	Activity specifics
Calorimetry:										
Direct	Single	Infant-elderly	VH	VH	VH	H–VH	H–VH	No	No	Yes
Indirect	Single-small	Young adult-elderly	H–VH	VH	VH	M–VH	H–VH	No	No	Yes
Job classification	Large	Employed only	L–M	L–M	L	L	L	Yes	Yes	?
Surveys:										
Indirect calorimetry diary	Single-small	Young adult-elderly	M–H	M–H	M–H	M–H	VH	No	No	Yes
Task-specific diary	Small-large	Adolescent-elderly	L–M	L–M	H–VH	VH	VH	?	Yes	Yes
Recall questionnaire	Small-large	Adolescent-elderly	L–M	L–M	M–H	M–H	L	Yes	Yes	Yes
Quantitative history	Small-large	Adolescent-elderly	L–M	L–M	L–M	L–M	L	Yes	Yes	Yes
Physiologic markers:										
Cardiorespiratory fitness	Small-large	Child-elderly	M–VH	M–H	M–H	M–VH	L	?	?	No
Doubly-labeled water	Single-small	Infant-elderly	H–VH	M–VH	M	M	L–H	Yes	Yes	No
Behavioral observation	Single-small	Infant-elderly	H–VH	H–VH	H–VH	L–H	L–VH	?	?	Yes
Mechanical and electronic monitors:										
Heart rate	Single-small	Infant-elderly	H–VH	M–VH	M–H	M–H	L–M	Yes	Yes	No
Stabilometers	Single-small	Infant	M–H	M	H–VH	L	L	Yes	Yes	No
Horizontal time monitor	Single-small	Child-elderly	M–H	M	H–VH	L–M	L–M	?	Yes	No
Pedometers	Single-large	Child-elderly	L–M	L	L	L	L–M	Yes	Yes	No
Gait assessment	Single-small	Child-elderly	H–VH	M–VH	L–M	M–H	L–M	?	Yes	No
Electronic motion sensor	Single-large	Child-elderly	M–H	L	L	L	L–M	Yes	Yes	No
Accelerometers	Single-large	Infant-elderly	L–M	L–M	L	L	L–M	Yes	Yes	No
Dietary measures	Large	Adolescent-elderly	M–H	M	M–H	M–H	L	Yes	Yes	No

[1] L = Low, M = moderate, H = high, VH = very high.

• To be nonreactive, the instrument must not alter the population or the behavior it seeks to measure.

Methods of Assessment

Calorimetry. In early energy balance studies, direct calorimetry measured energy expenditure through production of heat (*1*). With this assessment, physical activity is defined in relation to overall energy expenditure. Direct calorimetry has been used for many years and is highly accurate, with estimates of less than 1 percent error. Direct calorimetry, however, requires that individuals be sequestered in special chambers, making it expensive, limited to specific tasks, and impractical for the study of usual daily physical activities or large populations.

Indirect calorimetry measures the consumption of oxygen that closely correlates with heat production (*2–4*). The error is approximately 2–3 percent. Measurement of physical activity by means of indirect calorimetry requires that the participant wear a face mask or a mouthpiece with a nose clip, and a container for the collection of expired air.

The techniques of direct calorimetry and indirect calorimetry, although accurate, are not useful in epidemiologic studies because they alter or inhibit normal physical activity patterns and are prohibitively costly for use with large populations. However, these methods could be employed in samples of populations to validate more practical measures of activity.

Job classification. Job classification has been used to index physical activity by ranking jobs according to levels of activity and assuming that all persons in that occupational category expend similar levels of energy. This method, in use for many years, can be employed in large representative populations at minimal costs and is nonreactive. However, job classification has limitations for assessing physical activity and may not be sufficiently valid and reliable for epidemiologic use. Limitations include within-job classification variability, job intensity misclassification, secular changes in job requirements, seasonal changes in job requirements, possible selection bias, and omission of leisure and nonoccupational physical activity. These limitations may explain why epidemiologic studies using job classification schemes have not produced a clear pattern concerning the relationship of physical activity to coronary heart disease (*5*).

Variability of physical activity within a job class can be large. For example, cross-country truck drivers and local delivery truck drivers spend markedly different amounts of time sitting and unloading. This may contribute to a second problem: job intensity is often estimated differently by independent judges (*5–9*). Third, job requirements can change over time. For example, the extremely active occupation of the longshoremen in the 1950s now is almost completely mechanized and is a more sedentary occupation as the result of containerization (*9*). Fourth, the pattern of job-related physical activity may have seasonal variability. Postal carriers who must walk through winter snow expend much more energy than when walking during the summer. Fifth, a selection bias may exist as persons with illnesses transfer to less strenuous jobs. Hence, illness may be the cause of occupational activity rather than vice versa. Finally, job classification accounts for only the occupational subcategory of physical activity and therefore fails to identify the importance of leisure-time physical activity or those activities performed by the unemployed or retired (*10*). This may not have been an important omission in earlier epidemiologic studies because workweeks were long and leisure-time pursuits tended to exclude physical conditioning and vigorous sport play. Currently there is greater interest in measuring leisure-time physical activities because there is little variability in physical activity on the job in our highly mechanized society and potentially greater physical activity during leisure time.

Survey procedures. Survey procedures seek to acquire information from the participants about their physical activity and have four components (table 2). The first is the time frame the respondents are asked to remember. It may be as short as 5 minutes or as long as a year or more. A second component is the nature and detail of the physical activities. Participants may be asked to provide the frequency, duration, and intensity of specific activities or they may merely be asked if they have performed an activity or a group of activities. A third component is the mode of data collection. Personal interview, telephone interview, self-administration, mail surveys, or combinations of these are common methods. The remaining component is a summary index based on a calculated estimate of kilocalories expended or an ordinal scale that rank-orders persons according to their level of physical activity.

Based on these four characteristics, surveys may be grouped into four general types. Surveys that use short time frames (for example, less than 24-hour intervals) and are self-administered are commonly referred to as physical activity *diary surveys*. Surveys that obtain information about the past 1–7 days by means of a personal or telephone interview or a mail questionnaire are called physical activity *recall surveys*. Methods similar to the recall methods, yet which inquire about physical activities performed over a longer period, usually the past year, are often called *quantitative history* procedures. Surveys that solicit little specific information about the nature and detail of the physical activities are referred to as *general surveys,* regardless of the time frame of reference. The advantages and disadvantages of each type of survey are explained in the following sections.

'It should be evident that, while short-term diary and recall procedures result in highly detailed information, there is a limitation on the representativeness of the physical activity behavior or the population itself.'

Diary surveys. Although the diary technique has seldom been used in epidemiologic investigations of physical activity, it has been used in energy balance studies. For example, Edholm and colleagues used the results of a physical activity diary to estimate total daily caloric expenditure (*11*). The procedure they used first consisted of having individuals perform common tasks with concurrent measurement of energy expenditure by means of indirect calorimetry. Thereafter, individuals completed an ongoing diary, entering the specific tasks performed throughout the day. By summing the product of time spent in each task by the previously measured rate of energy expenditure for each activity, an overall estimate of total daily caloric expenditure was made. A direct comparison of the results of the ongoing diary with other estimates of total daily energy expenditure, measured either through indirect calorimetry or through caloric intake, revealed that diaries could be highly accurate indices of daily energy expenditure (*11,12*). However, although the technique is very accurate, it suffers from cost, time, and acceptability constraints. In addition, persons may be unwilling to record every physical activity they do throughout the day, or they may alter their normal pattern of physical activity to simplify the recording process. Each of these limitations makes this procedure of little value for

Table 2. Characteristics of various physical activity survey procedures used in epidemiologic investigations

Type of survey	*Time frame of survey*	*Nature and detail of activity*	*Time unit or level of detail*	*Mode of collection*[1]	*Summary index of physical activity*
Diary:					
Edholm and coauthors (*11*)	24 hours	Specific activities	Activity entered for each 24-hour period for 2 weeks	SAQ	Kcal score
LaPorte and coauthors (*16*)	24 hours	Specific activities	Activity entered for each 4-hour period for 12 hours	SAQ	Kcal score
Recall:					
Baranowski and coauthors (*38*)	1 day	Specific aerobic activities	Probed for 20 minutes of aerobic activity	SAQ	Kcal score
Morris and coauthors (*17*)	2-day total (Fri. and Sat.)	Specific activities	Activity for each 5-minute segment	MQ, SAQ	Grades I, II, III, IV, VE (vigorous exercise)
Bouchard and coauthors (*21*)	3-day total (2 week-days)	Intensity classes	15-minute intensity classes	SAQ	Kcal score
Paffenbarger and coauthors (*22*)	Past week	Stairs + blocks per day	Total flights, blocks	MQ, SAQ	Kcal score
		Sports, recreation	Time per week		
		Intensity classes .	Hours in each class		
Stanford (*23,24*)	7 days	Intensity classes .	Hours per week in each class	PI	Kcal score
		Walking activities	Yes or no		
		Aerobic activities	Yes or no		
Quantitative history:					
Montoye (*26*)	Year	Specific activities in work and leisure	Time per occasion, months per year, occasions per month	PI	Kcal score
Taylor and coauthors (*25*)	Year	Specific activities in leisure	Time per occasion, months per year, occasions per month	PI	Kcal score
General:					
Shapiro and coauthors (*28*)	1 day	General work and leisure	General statements of participation level	MQ, SAQ	4 classes for work and leisure
Framingham (*29*)	1 day	4 intensity classes for work and leisure	Hours per day in each class	PI	Ordinal scale, score of 24 or higher
Salonen and coauthors (*30*)	1 week	4 intensity levels for work and leisure	General statements of level of intensity	SAQ, PI	Assigned 1 of 4 levels for work and leisure
Magnus and coauthors (*32*)	1 year and past 2 years	Walk, cycle, garden activity (WCG)	Hours per week, months per year	PI	WCG groups: occasional, seasonal, habitual; at 1–4, 4–7, or 7+ hours per week
		More strenuous than WCG	Yes or no		
Baecke and coauthors (*33*)	1 year	Work, leisure sport and nonsport	5-point scale for each question	MQ, SAQ	Ordinal scale
		Sports participation	Hours per week, months per year		
		Walking and cycling	Minutes per day		
Morrison and coauthors (*34*)	1 week	8 leisure activity intensity classes	0, 1, 2, 3, 4+ times per week for 30 minutes	PI	Assigned 1 of 8 classes of leisure activity
Lipid Research Clinics (*35*)	1 week	General work and leisure	2 general questions of strenuous and regular labor or exercise	PI	Classed inactive, moderately, active, or very active

[1] PI = Personal interview, MQ = mail questionnaire, SAQ = self-assessment questionnaire.

epidemiologic studies; nonetheless, it represents an accurate standard to approximate.

The technique of Edholm and colleagues has been modified to eliminate the need for actual assessments of task-specific energy expenditure. These modified techniques use published values available in sources such as Passmore and Durnin (*13*) and use the method of scoring kilocalories employed in the study by Edholm and colleagues (*11*). It should be noted, however, that using published values may produce variability in estimating energy expenditure because there may be relatively large differences between individuals for certain tasks (*14*). Moreover, the published values typically have been produced through the assessment of physical activity for young, college-age males who are often atypical of the populations generally measured in epidemiologic investigations of chronic diseases, such as women, blacks, children, and obese persons. Although diary procedures using published values of activity intensity may not provide accurate estimates of caloric expenditure, they seem to be adequate to rank-order individuals according to overall activity levels (*15*).

Another modification of the Edholm diary technique had persons record the physical activities they performed during 4-hour periods ending at 12 noon, 4 p.m., and 8 p.m., rather than record continuously throughout the day (*16*). Such a method may make the diary procedure less arduous; however, the likelihood of forgetting important details of physical activity participation may increase.

'As evidenced by this review, there are more than 30 different techniques that have been employed for assessing physical activity in population studies. Perhaps with the exception of dietary and nutritional assessments, we know of no other health-related behavior that has been measured in so many ways.'

Recall surveys. The diary technique of recording physical activity participation throughout the day is relatively difficult to employ on a population basis because of the high demands placed on the participant. More often, epidemiologic research of physical activity has used the recall procedure. The recall procedure used by Morris and coworkers (*17*) was derived from a questionnaire-interview procedure developed by Yasin (*18*). Yasin pretested the questionnaire using 140 male British civil servants administered personal interviews four times during a year. Time spent in reported activities was converted to kilocalories, using published intensity values. There was a high correlation between the questionnaire and the quarterly assessments, suggesting that the survey could index overall physical activity. Additionally, the rankings appeared to be associated with caloric intake, with persons in the highest third reporting a higher caloric intake than those in the lowest third (*19*). Reliability and validity of this instrument therefore appeared to be good. Morris and colleagues modified Yasin's interview so that it could be self-administered. The recall questionnaire consisted of a 5-minute by 5-minute recall of a Friday and Saturday. Marr and colleagues compared the kilocalorie score of Morris's survey in a small group of these men against weighed dietary intake and found a correlation of .40 (*20*). Hence, an indirect validation seems to yield favorable results. The reliability of Morris's questionnaire was untested but assumed to be similar to that of Yasin.

Bouchard and colleagues employed another daily recall in a study of 150 men and 150 women who were 10–50 years old (*21*). The subjects provided a 15-minute by 15-minute recall for 3 days (2 weekdays, 1 weekend day). The activity for each segment was to be scored from 1 to 9, using published values for oxygen consumption, and was thereafter converted to a kilocalorie score. Repeat administration yielded a reliability correlation of $r = 0.96$. The highest intensity category was most reliably repeated. In addition, Bouchard and coworkers indirectly validated this recall against physical work capacity on a bicycle ergometer and with body fat estimates and found a small but significant relationship (*21*).

A recall survey by Paffenbarger and coauthors probed for the distance walked, stairs climbed, and sports or recreational activities undertaken during the previous week (*22*). As such, this survey attempted to ask about the major contributors to the weekly physical activity pattern. The questionnaire is relatively simple to implement through the mail but has not been tested for reliability or validity.

A somewhat different recall procedure has been developed at Stanford University that has promise in the assessment of physical activity. In this technique individuals recall the time spent doing activities of several levels of intensity over the previous 7 days rather than provide detailed time esti-

'To obtain the best assessment of physical activity that does not bias the sample, epidemiologists therefore need to balance increased accuracy of these instruments with representativeness of both the behavior and the population.'

mates for specific activities (*23*). Thus, the measurements are thought to be a more accurate assessment of intensity-related issues. This 7-day recall has been indirectly validated using results from a community health survey, a randomized clinical trial, and two worksite health promotion programs yielding expected relationships with cross-sectional physical activity patterns of men and women; changes in maximal oxygen consumption, body fatness, high-density lipoprotein (HDL)-cholesterol and triglycerides; as well as increased physical activity over time due to intervention efforts (unpublished observations, Steven N. Blair, 1985). A short-term repeatability of $r = 0.67$ has been established (*24*). Future research employing this technique is likely to prove valuable as a physical activity assessment tool.

Quantitative history surveys. These surveys have time frames greater than 1 week and request detailed information on specific activities. In the questionnaires of Taylor and coworkers and Montoye, persons recall over the previous year their activity patterns for a list of specific physical activities (*25,26*). The Montoye questionnaire had been indirectly validated by comparing results of the objective summary kilocalorie index against subjectively ranked estimates of physical activity level made by an independent judge, yielding high correlations of .86–.97. Further, the questionnaire was compared against rank-ordering of caloric intake, which differentiated the 20 percent least active and 20 percent most active. The middle 60 percent were less well discriminated from the 20 percent most active. Further, indirect validation showed relationships with body fat estimates and blood lipids (*10*).

The Taylor questionnaire has been indirectly validated against treadmill work performance in 175 men and with estimated caloric intake in a comparison of female college swimmers and college students (*25*). Further, relationships were found between results of the Taylor questionnaire and HDL cholesterol, body mass index, and heart rate (*27*). Short-term repeatability (roughly 2 weeks) of the Taylor questionnaire was recently examined in a group of middle-age men having previous experience with the questionnaire. Results suggest that the Taylor questionnaire has a correlation between first and second report of $r = 0.85$ for total leisure and 0.73–0.79 for activity groupings of light, moderate, and heavy intensity (unpublished observations, C. J. Caspersen, 1983). Continued breakdowns into activity class categories (such as sports, conditioning, lawn and garden maintenance, and home repair) or the individual activities themselves showed lower repeatability because of a greater number of parameters to be remembered correctly. Overall, the quantitative history surveys can be implemented on a population basis, yielding enormous detail on the physical activity pattern.

General surveys. Another type of assessment that has been used is the general physical activity survey. Surveys that fit into this category provide less detail than other techniques. In fact, the participant may merely provide a subjective impression of his or her usual physical activity or otherwise may simply select one of several descriptive classes.

One example of a general survey, which was developed by Shapiro and colleagues for the New York Health Insurance Plan, had separate sections for work and leisure-time physical activity (*28*). The questionnaire was mailed to plan members and could be easily self-administered in about 5–10 minutes. Scored results of the survey could range from 1 to 28 for work-related physical activity and 0 to 10 for the leisure-time physical activity. Cutoff points were determined such that both sets of scores could be divided into four classes. In an effort to determine the consistency of the self-report, a followup interview of 38 persons was conducted. Despite the survey's simplicity, those persons in the followup group were repeatedly placed in the same work, leisure, and total work and leisure classes only 55–61 percent of the time. There has been no reported attempt at validation of this questionnaire.

The Framingham Study of heart disease also employed a general physical activity survey for personal interviews that queried the hours per day spent in sleep, work, and extracurricular activities (*29*). For the latter two categories, the hours spent in sedentary, slight, moderate, and heavy activity were probed for and multiplied by weighting factors of 1.1, 1.5, 2.4, and 5.0, respectively. Sleep hours were multiplied by a weighting factor of 1.0. The weighting factors were chosen to parallel the increased rate of oxygen consumption associated with increasingly more intense physical activity. The lowest possible summary score of this survey was 24 (for example, 24 hours of sleep). Although this survey was capable of rank-ordering individuals by physical activity level, there are no published reports on the repeatability or validity of this index.

Salonen and others of the North Karelia Study employed four alternative questions to index work and leisure physical activity separately (*30*). The work alternatives ranged from a category of work activity described as mostly sedentary to a category of regular, hard physical labor such as digging, lifting, and carrying burdens daily. Similarly, the leisure-time physical activity alternatives ranged from a category of almost complete inactivity to a category of regular hard physical training, such as running and soccer, several times per week. No repeatability of this survey has been reported. However, Grimby and coworkers used essentially the same questions to place 641 men into groups of the two lowest and two highest categories for both work and leisure questions. The investigators found a significant relationship with predicted maximal oxygen consumption, thereby achieving a simple level of indirect validation (*31*).

Another general survey used in the Netherlands by Magnus, Matroos, and Strackee asked questions to determine the total hours per week spent in walking, cycling, and gardening activities as a group (*32*). Further, subjects were queried about three categories of yearly participation for these combined activities: habitual (participate more than 8 months a year), seasonal (4–8 months), and occasional (less than 4 months). Another question asked about the time spent in low-level "moving about" activities; sedentary status was ascribed to persons who performed such activities 2 hours per day or less. Although this survey is quite useful for sorting out issues of activity seasonality there have been no published reports of repeatability of this survey or of validation attempts.

Another general survey used in the Netherlands was developed by Baecke and colleagues (*33*). In its original form the questionnaire comprised 29 items. However, after performing a principal-component analysis, 16 items representing three meaningful factors emerged: physical activity at work, sport activity, and nonsport leisure-time physical activity. Questions relating to specific sport participation and occupation were open-ended; the remaining questions were based on five-point scales for subjective ranking. The scored questionnaire provided ordinal values ranging from a minimum of 16 for a sedentary person to approximately 70 for a person

'The recall, general, and quantitative history survey procedures currently offer the best compromise as epidemiologic measures of physical activity in large populations.'

with considerable work, sport, and nonsport leisure-time physical activity. This self-administered questionnaire was mailed to 306 men and women, ages 19–31 years, and was followed by a debriefing at a clinic interview. The same questionnaire was readministered during a home interview. Reliability coefficients were acceptable, being between $r = 0.81$ and $r = 0.88$ for work and sport and $r = 0.74$ for nonsport leisure. Indirect validation was performed by correlating results of the questionnaire with education, lean body mass, and subjective experience of workload (SEWL). The SEWL was a Dutch series of 57 yes-or-no questions about the perceived physical stress of many common tasks. Baecke and coworkers found for both men and women that work activity was inversely associated with educational level; that leisure nonsport activity was positively associated with educational level; and that there was an inverse association between SEWL and both leisure sport and nonsport physical activity. For women, a positive association between leisure sport and educational level was found; for men, an association was found between both work and leisure sport activity and lean body mass. In all, the questionnaire seemed to be both reliable and at least indirectly validated.

Morrison and colleagues of South Africa recently developed a general physical activity survey that assigns individuals to one of eight classes of exercise training activities (*34*). Persons reporting no exercise training at all were assigned to class one. Persons reporting one, two, and three or more occasions per week of low-intensity training in activities such as golf, bowling, "keep-fit" activities, yoga, or 30 minutes of walking were assigned to classes two through four. Persons reporting one, two, three, and four or more times per week of high-intensity activities, such as soccer, tennis, or 30 minutes of walking, were assigned to classes five through eight. There are no published reports regarding repeatability of this classification; however, an indirect validation was performed, using estimated maximal oxygen consumption, and it yielded a correlation of $r = 0.67$.

In addition to the general surveys just described, several attempts at overall ratings of physical activity have been attempted. The Lipid Research Clinics Program Prevalence Study employed two questions regarding the regularity and perceived intensity of strenuous work or leisure activity (*35*). These simple questions appeared to index individuals accurately into three categories that were in turn related to physiological parameters. Similarly, the National Health Interview Survey had individuals simply rate their physical activity level with respect to persons their own age (*36*); the results were successfully used in rank-ordering populations. Still another type of overall assessment pertains to the assessment of functional ability. Surveys of disability, impaired movement, and activities of daily living are designed to assess physical activity levels that are below those of able-bodied populations. These surveys index the ability of individuals to perform certain movements or daily activities such as walking, climbing stairs, and bathing. The assumption is that persons with the greatest impairment are also the least active. These surveys may be very important for indexing persons at the lower end of the physical activity spectrum—the 15 percent of the U.S. population with impaired mobility (*37*). However, details are generally lacking regarding the reliability and validity of any of these instruments.

The recall, general, and quantitative history surveys have been increasingly used in epidemiologic investigations of physical activity because of their ease of implementation. Currently, they are the most practical means of measuring physical activity in large populations. There are certain issues that need to be considered, however, with these procedures. For example, the degree to which these instruments are accurate across populations is not known (for example, plowing requires quite a different expenditure of energy in Mexico than in Iowa). Assessment of physical activity cross-culturally, such as black populations compared with whites ones or even women compared with men, may not reflect differences in activity but rather in the perception of how to complete the surveys.

Further, each survey procedure has one significant limitation: the capacity of a person to remember details of past physical activity. At present, few studies have addressed this issue; however, one study is particularly noteworthy. Baranowski and colleagues examined some of the limitations of a 24-hour recall study of children by comparing results with direct observation throughout the day (*38*). Although the study only asked about aerobic-type activity, the daily recall seemed to be improved when context-specific times, such as before school, during school, and after school, were probed for. The study clearly points to the need to study methods of improving memory when developing physical activity surveys.

It should be evident that, while short-term diary and recall procedures result in highly detailed information, there is a limitation on the representativeness of the physical activity behavior or the population itself. Generally, short-term physical activity surveys encompassing 1–7 days are more likely to miss the physical activity behavior not performed during other seasons. One way to overcome this limitation is to readminister the survey during other seasons and then to pool the summary indices. Unfortunately, subjects may be unwilling—or, more importantly, unavailable—to make repeated assessments. Further, repeated surveys require more effort by the investigators. On the other hand, the quantitative physical activity history is capable of covering seasonal variation but at the expense of the capacity of the individual to remember over long periods. Further, as the demands on the subject to recount minute details increase, both ability and willingness to comply decrease.

As noted, the survey procedures each derive a summary index capable of rank-ordering persons by their level of physical activity. In other instances it is desirable to identify participation in more specific activities or classes of activities. For example, the Stanford 7-day physical activity recall employs a series of questions relating to walking and aerobic activity participation because each may be related to different health outcomes (*24*). Hence, this survey both provides an overall kilocalorie index and, at the same time, specifically details involvement in common high- and low-level physical activity. Other surveys may need to use similar questions to pinpoint more specific activity participation.

Physiological markers of physical activity. The fact that changes in the level of vigorous physical activity are known to influence cardiorespiratory endurance has led to the frequent use of maximum oxygen consumption to estimate physical activity. For small, nonpopulation-based samples the correlation between reported physical activity level and physical work capacity is modest (*25*); however, for large, population-based studies the relationship is weak (*39*). This may be because a large portion of cardiorespiratory endurance is genetically determined (*40,41*). Therefore, cardiorespiratory endurance may be a relatively poor index of physical activity (*42*). It should be noted that, for physical activity assessments that provide summary scores highly dependent on the intensity of physical activity participation, cardiorespiratory endurance may serve as a suitable indirect validation criterion.

Another physiological approach, the doubly labeled water technique, appears to provide an integrated measure of energy expenditure over time. In this method subjects ingest water containing isotopically labeled hydrogen and oxygen atoms of negligible health risk to the participants. Measurement of the relative proportions of unmetabolized water and water that has been incorporated into the energy cycle can provide an overall estimate of energy expenditure (*43,44*).

Originally this method was economical only in studies of small animals. However, declining costs for the isotopes, along with technological advances in isotopic mass spectrometers, have made it economically feasible to apply this procedure in studies involving humans. The method appears to be quite accurate, with error estimates ranging from 2 to 10 percent when compared with calories from weighed dietary intake (*43–46*).

In addition, the doubly labeled water technique can be used with a person of any age, does not restrict free-living physical activity, requires minimal cooperation by the subject, and is generally acceptable to the subject. The technique takes a minimum of 2–3 days, and can extend through several weeks. Although this technique has had encouraging results, the cost for the isotopes is still about $225 per person, making it prohibitive for large population studies but perhaps most useful for validation studies. In addition, the technique provides only an overall kilocalorie index that cannot identify specific types or patterns of physical activity participation that may be very important to assess.

Behavioral observation. Observational techniques have been developed by behaviorists as an approach to the monitoring of physical activity. In one approach, an observer watches an individual and rates that observed activity at specific time intervals (*47,48*). The ratings are used as an estimate of the physical activity level. Random photographs and judgments of activity have also been employed (*49*). The assessment of activity through the use of observational techniques has inherent appeal; however, it is impractical on a population basis. Moreover, it is likely that only a select group of volunteers would consent to being continually observed. Also, the observations themselves may influence typical behavior. Thus, the assessment of activity through behavioral sampling is likely to have only limited applicability on a population basis. The real importance of this technique, like the doubly labeled water technique, may be its usefulness as a validation criterion for certain types of measurement instruments (*38*).

Mechanical and electronic monitoring. A number of mechanical and electronic instruments have been developed to assess body movement or heart rate responses to physical activity. Although some have been around for many years, technological advances continue to make them viable methods of objective assessment.

Heart rate monitoring. Recent advances in heart rate monitoring have made it feasible to obtain continuous heart rate recordings over an extended time. Several reports in normal populations have assessed physical activity through the use of heart rate monitoring approaches (*50–52*). Heart rate monitoring is attractive because it directly measures a physiological parameter known to be related to physical activity and because it provides a continuous record that may reflect both intensity and duration of daily activity.

To estimate energy expenditure, one must assume a close linear relationship between heart rate and oxygen consumption. Heart rate and oxygen consumption are typically measured for each subject over a range of work rates on a cycle ergometer, treadmill, or step bench. Regression curves for the heart rate and oxygen consumption relationship are developed for each person (*52,53*).

These initial laboratory measures are time-consuming and expensive and may eliminate many subjects, thereby biasing the sample; thus, this technique is impractical for most epidemiologic research. In addition, the assumption of a linear heart rate and oxygen consumption relationship has also been shown to be affected by the amount of muscle mass involved in the activity (*54*), the type of muscular contraction (*55*), environmental temperature (*56*), state of physical training (*57*), fatigue (*58*), and emotional stress (*59*).

Given the potential number of confounding factors affecting heart rate, it is not surprising that in the few available reports where the daily energy expenditure has been assessed using the heart rate method the results are rather poor (*60*). It may be possible to avoid the problem of subject calibration and obtain information on physical activity from heart rate monitoring by considering the difference between resting heart rate and mean daily heart rate (*61,62*). This would make the use of heart rate monitoring more practical for epidemiologic research and warrants further evaluation.

Motion sensors. The assessment of physical activity by measuring "movement" is appealing because more active people typically move more than less active people. Also, movement may measure physical activity more accurately than estimates of energy expenditure. Clearly there are many factors that influence energy expenditure and the physiological measurement of energy expenditure such as specific dynamic action, basal metabolic rate, and perhaps body weight, ambient temperature, and age (*63*). Movement per se, however, may be highly complex itself and require indepth analysis of type, frequency, and intensity to determine its physiological effects.

Researchers in the 1960s and 1970s developed several excellent approaches for measuring movement in infants using stabilometers on mattresses where continuous 24-hour recording could be easily done (*64,65*). The measures appear to be relatively accurate assessments of activity in infants but of no utility for children or adults (*66*).

Another instrument, affixed to the subject's thigh, has been used to compare the amount of time spent seated or lying down with that spent standing or moving about (*67*). There is no information concerning either the reliability of the instruments or the practicality of using the instruments with large populations.

The primary type of measurement of body movement has been through the use of pedometers (*68*). Pedometers are instruments designed specifically to evaluate walking behavior, and several different types are available commercially. Reports evaluating the accuracy of the instruments have identified inter- and intra-instrument variability primarily associated with the mechanical fulcrum of the instruments, each responding inconsistently to a given force (*69,70*). There is little information concerning the applicability of pedometers across diverse populations and whether the instruments can index levels of individual and group activity.

Instruments from gait research have not as yet been employed in population studies but may afford a major increase in accuracy for assessing walking behavior. These instruments fit into the shoe and measure not only the step frequency, as with a pedometer, but also the force applied with each step (*71*). Marsden and Montgomery (*72*) found that in-shoe step counters could easily discriminate activity between postal carriers and school children, as well as different activity patterns by individuals. There has been little followup on the use of in-shoe step counters in population studies, and their reliability and validity are not established. These instruments have not been used in population studies but may prove to be valuable in epidemiologic studies of osteoporotic women where weight-bearing activity appears to be critical to bone strength (*73*).

The physical activity monitors that have received the most interest are electronic motion sensors, now commercially available (*74–77*). Electronic motion sensors can be classified into instruments that only assess the quantity of movement and instruments that assess both the quantity and intensity of movement (accelerometers).

The large-scale integrated motor activity monitor (LSI) is an instrument about the size of a pocket watch. It measures body movements and can be placed on various body locations. Within the monitor is a cylinder with a ball of mercury. A 3 percent inclination or declination of instruments closes a mercury switch, which registers in an internal counter. The results of the counter can be read by holding a magnet to the side of the instrument, thereby activating a light-emitting diode (LED) display indicating the number of movements (*16*).

LSI units have been shown to have low variability between units (*78*). Hence, it can safely be concluded that one LSI is measuring the same as another unit and that the instruments appear to measure correctly across time. Moreover, the LSI units have been able to discriminate between varying population groups (*8,79*).

Accelerometers, which measure both frequency and intensity of movement may prove even more helpful than LSI units and have recently been introduced commercially (*77,80*). Future methodologic studies need to be completed to evaluate the effectiveness of this instrument for population study use.

The mechanical and electronic monitors of heart rate and movement may be highly useful in assessing physical activity, at least in small groups. Their appeal is that they require little time from the investigator as well as limited effort by the subject. Further, many monitors do not interfere with or influence physical activity and seem to be personally and socially acceptable. Currently the cost of these devices makes large population studies unlikely; they provide information on specific categories or types of physical activity, but not an estimate of energy expenditure. As technology improves and the costs decrease, these monitors may be made applicable to population studies of physical activity.

Table 3. Correlations between activity monitor readings and surveys

Study	Population	Activity monitor component	Survey technique	Pearson correlation	Probability
LaPorte and coworkers (*16*)	20 male college students	Total daily activity	Diary	.69	< .01
		Total daily activity	Taylor questionnaire	.05	> .05
LaPorte and coworkers (*87*)	42 male industrial workers	Total daily activity	Taylor questionnaire	.02	> .05
		Total daily activity (workday)	Taylor questionnaire	.11	> .05
		Total daily activity (off day)	Taylor questionnaire	.45	< .05
	79 menopausal females	Total daily activity	Paffenbarger questionnaire	.23	< .05
Cauley and coworkers (*88*)	256 menopausal females	Daytime activity	Paffenbarger questionnaire	.21	< .05
		Total daily activity	Paffenbarger questionnaire	.13	< .05

Dietary measures. The caloric value of food intake may be used as an estimate of energy expenditure and, hence, of physical activity, if one assumes that energy balance has been achieved with stable body weight. Edholm and colleagues demonstrated that long-term assessment of total caloric intake, using either direct analysis of equivalent food samples or published values, provides a highly accurate assessment of caloric expenditure for an individual (*11*).

However, it is well known that total caloric intake is influenced by the person's physical activity level and also by the total body weight of the subject. For example, two persons having markedly different body weights may have the same total daily caloric expenditure with the lighter person doing more physical activity than the heavier person. Further, using weighed dietary intake as a measure of physical activity is costly and requires great demands on the subject.

To address these issues, Sopko and colleagues recently analyzed data from two carefully controlled feeding experiments, one of obese subjects and the other of normal-weight subjects, using precisely weighed dietary samples for total daily caloric intake (*81*). After adjusting the total caloric intake by the individual's body weight in kilograms (kcal per kg per day), the investigators were able to show a strong negative correlation with body fat ($r = -0.79$) and an equally strong positive association with maximal oxygen consumption ($r = 0.76$), each serving as an indirect validation of physical activity level. In addition, Sopko and coworkers employed a 3-day food recall to estimate daily caloric intake for their obese subjects. When the investigators compared total caloric intake in kilocalories per kilogram per day with maximal oxygen consumption and body fatness, weaker yet still significant correlations of $r = -0.54$ and $r = 0.31$, respectively, were noted.

It is important to recognize that estimates of dietary intake are also known to have considerable variability, as shown by the smaller correlations found by Sopko and coworkers when the 3-day food recall was used. Beaton and colleagues demonstrated that estimated total caloric intake in a 24-hour dietary recall has considerable intra- and inter-individual variability (*82*) and may therefore be inadequate for indexing an individual's physical activity level within a group. Hence, dietary measures may have to improve considerably to be a useful and practical index of physical activity. Further, dietary measures of physical activity are unable to identify the types, frequency, intensity, or duration of physical activities.

Relationships Between Measures

Researchers have begun to evaluate the relationship between measurement techniques to determine the degree to which each technique measures the same thing. For example, Buskirk and coworkers compared results of the Health Insurance Program and Montoye questionnaires and noted significant correlations of relatively low magnitude (*83*). This indicated that the two instruments, each attempting to measure work and leisure-time physical activity, were less than congruent. In addition, the LSI motion sensor has been compared with various survey techniques in several populations having correlations of .70 with diaries in several populations; however, correlations with recall procedures were much smaller (table 3).

'Nonetheless, despite these crude and varied tools a relatively consistent pattern has been shown between increased physical activity and reduced risk of coronary heart disease, osteoporosis, and noninsulin-dependent diabetes.'

Further, it has recently been suggested that the Paffenbarger survey (*22*) appears to be related to HDL_3-cholesterol whereas activity monitoring appears to be associated with HDL_2-cholesterol and bone density (*79*). Thus, the component of physical activity measured by these two instruments may not be the same. Moreover, each may be associated with a specific aspect of health.

Discussion

As evidenced by this review, there are more than 30 different techniques that have been employed for assessing physical activity in population studies. Perhaps with the exception of dietary and nutritional assessments, we know of no other health-related behavior that has been measured in so many ways.

Part of the reason for the large number of physical activity measures is that there are many intercorrelated dimensions of physical activity, and each may be associated with a specific aspect of health. As examples, the absolute caloric expenditure regardless of intensity is associated with obesity; weight-bearing activity, regardless of the aerobic nature of the specific task, appears to be associated with bone density; and intense aerobic activity is associated with cardiorespiratory fitness. Hence, the specific dimension measured would depend on one's interest in specific conditions such as osteoporosis, noninsulin-dependent diabetes, or obesity, thereby varying in importance from study to study.

For each physical activity assessment employed, a different operational definition exists. In the most general sense, physical activity is defined by the level of caloric expenditure (*84*). While this has been useful in studies of obesity and coronary heart disease, it is somewhat like defining diet simply as caloric intake. Evidence is beginning to emerge that specific types and patterns of physical activity, rather than absolute levels of physical activity, may have differential associations to health (*85*). It is important that we identify these associations. However, regardless of the dimension of interest, the most appropriate operational definition and measurement method must be established.

Upon choosing the most appropriate measure of physical activity it is important to establish instrument reliability and validity to ensure accurate assessment. Several studies have begun to determine the reliability of instruments. At minimum, reliability has been estimated by test-retest reliability coefficients; intraclass correlations have yielded considerably greater detail (*21*). The validity of the measurement instrument is even more critical but poses serious problems because there is no accepted validation criterion. Various instruments have used caloric intake, direct observation, and physical work capacity as validation criteria; however, the results have been unclear because the definition of physical activity has typically been absent in the study or inconsistent between instruments. Clearly, the validation criterion should depend on the specific operational definition of physical activity. For example, as validation criteria the following might be considered: surveys that derive kilocalorie scores may wish to use doubly labeled water; surveys concerned with intense aerobic activity might reasonably employ maximum oxygen uptake measures; motion sensors might use behavioral observation; and walking surveys might employ a pedometer or an in-shoe step counter. Hence, the validation criterion may vary depending on the dimension and definition of activity used in a measurement instrument.

It is also important to determine the sources of variability in physical activity data. Beaton and coworkers recently revealed that variability of 24-hour dietary recall can be partitioned into intra- and inter-individual, day of the week, sequence of administration, interviewer, and methodological components (*82*). It was discovered that information from these components determined the sampling techniques and replications necessary for representative dietary data. Such issues are not well understood in physical activity research but are particularly urgent in physical activity survey procedures to index physical activity adequately on individual and group levels.

Another important source of variability in adequately characterizing physical activity is the accuracy of ascribing intensity. The intensity of activities such as walking, jogging, and running clearly vary with velocity, the surface being covered, and whether hills are encountered; each variable, alone or with the others, results in markedly different caloric expenditure and physiologic outcomes. However, only one intensity factor for each activity has generally been used (*25*). Because walking is so prevalent, it may be necessary to identify ways of improving intensity estimates for the varying levels of walking participation.

Ideally one would like the most accurate assessment of activity. However, increasing the accuracy of the instruments typically involves increased demands on subjects, resulting in limitations in behavior or participation in the measurement. For example, the accuracy of direct and indirect calorimetry and of direct observation is quite high; however, the demands on the subject are even higher. To obtain the best assessment of physical activity that does not bias the sample, epidemiologists therefore need to balance increased accuracy of these instruments with representativeness of both the behavior and the population.

The recall, general, and quantitative history survey procedures currently offer the best compromise as epidemiologic measures of physical activity in large populations. There is little question that these survey procedures are more precise than job classification and can be used on equally large populations. The procedures appear to be relatively reliable and unlikely to alter normal daily physical activity; they do not produce major selection bias and are inexpensive to administer. Unfortunately, they lack the objectivity of electronic and mechanical monitoring. The principal problems with recall procedures are that little is known about the dimensions of the physical activity being measured, that recall relies on the participant's cooperation, and that the reliability and validity of recall are often incomplete or undetermined.

Future developments in electronic monitoring are likely to reduce cost and increase precision, so that they may be practical for large-scale population studies. Although they are only in the experimental stage, they appear to be highly reliable. Unfortunately, we know little of what these monitors actually measure or what effect monitoring has on population selection biases.

The multitude of instruments and the limited knowledge about what is being measured at first approximation appear discouraging for epidemiologists interested in physical activity. Nonetheless, despite these crude and varied tools, a relatively consistent pattern has been shown between increased physical activity and reduced risk of coronary heart disease, osteoporosis, and noninsulin-dependent diabetes (*86*). By identifying the most important dimensions of physical activity and with improvements in their assessment, relationships of physical activity to disease and health will likely be stronger and more consistent.

Conclusions

More than 30 different techniques are available for assessing physical activity, with many different variations in the techniques. Most may be grouped as calorimetry (direct and indirect), job classification, survey procedures, behavioral observation, electronic and mechanical monitors, physiological markers, and dietary measures. These range from highly precise assessments to crude measures and are based on very different definitions of physical activity—from work activity to leisure-time activity to cardiorespiratory fitness, movement, and heart rate. It is clear that "physical activity" measured in some studies is not the same type of "physical activity" measured in other studies. At present, a number of relationships have been recognized between several of these measures and indices such as body fat, cardiorespiratory endurance, and HDL-cholesterol. Although these different assessment procedures may not be highly related and may at times be crude or inaccurate, we are beginning to see patterns or relationships between them and indices of health and disease.

The following points indicate what is currently known or strongly suspected about physical activity assessment:

- Physical activity is a complex behavior having many interrelated dimensions.
- Physical activity measures may be classified in at least seven ways: calorimetry, job classification, survey procedures, physiological markers, behavioral observation, mechanical and electronic monitors, and indirect dietary estimates.
- Each class of physical activity measure captures only a part of the entire activity behavior pattern. For example, some instruments can assess energy expenditure; others, the frequency, intensity, duration, and type of movements; and still others, weight-bearing physical activities.
- Some dimensions of physical activity are related to health or disease outcomes or states. For example, energy expenditure is related to obesity; high-intensity aerobic activity is related to cardiorespiratory endurance; and weight-bearing physical activity is related to osteoporosis.
- Research and practical considerations govern the selection of physical activity measures. The research issues pertain to survey design and popula-

tion, potential to interfere, and specific interest in dimensions of health, disease, or both outcomes or states. Practical issues pertain to cost, time, and the measurement acceptability.

• Survey procedures are currently the most practical physical activity measure for large-scale population studies.

• There are at least four characteristics of survey procedures: the time frame respondents are asked to remember; the nature and detail of the physical activities; the mode of data collection; and the summary index derived by survey results to rank-order individual physical activity levels.

• The motion sensors may prove to be extremely useful for large-scale population studies of physical activity if their cost can be reduced and their validity determined.

• Despite the difficulty in measurement a relatively strong association has been found between physical activity and health, suggesting that with improvements in assessment techniques even stronger associations should be seen.

Recommendations

The following specific recommendations have been suggested as a result of what we need to know regarding the assessment of physical activity.

1. Continue to explore the relationships between the different dimensions of physical activity with specific health and disease outcomes or states.

2. Determine the best methods for measuring the different dimensions of physical activity and for the outcomes or states of interest.

3. Establish better estimates of reliability and validity for each physical activity measure, both for the entire population and for population subgroups.

4. Determine the most appropriate measurement technique for population subgroups such as women, elderly, children, cultural groups, and ethnic groups. Most of our current knowledge of physical activity assessment is based upon the study of white, middle-aged men.

5. Determine the factors that enhance the recall of physical activity in survey procedures for the entire population and for population subgroups (such as children, the elderly, and various cultural and ethnic groups).

6. Determine more accurate methods to estimate the level of intensity of physical activity participation in survey procedures.

7. Determine a minimum set of standardized survey questions that can be included in all instruments, regardless of the specific focus of the physical activity assessment, in order to compare all populations uniformly. To such a core of questions others could be added as needed.

References

1. Passmore, R.: The regulation of body weight in man. Proc Nutr Soc 30: 122–127 (1971).
2. Orsini, D., and Passmore, R.: The energy expended carrying loads up and down stairs: experiments using the Kofranyi-Michaelis calorimeter. J Physiol 115: 95–100 (1951).
3. Dauncey, M. J., Murgatroyd, P. R., and Cole, T. J.: A human calorimeter for direct and indirect measurement of 24-h energy expenditure. Br J Nutr 39: 557–566 (1978).
4. Harrison, M. H., Brown, G. A., and Belyavin, A. J.: The "Oxylog": an evaluation. Ergonomics 25: 809–820 (1982).
5. Leon, A. S., and Blackburn, H.: The relationship of physical activity to coronary heart disease and life expectancy. Ann Ny Acad Sci 301: 561–578 (1977).
6. Gabarant, D. H., Peters, J. M., Mack, T. M., and Bernstein, L.: Job activity and colon cancer risk. Am J Epidemiol 119: 1005–1014 (1984).
7. Davis, H. L., Faulkner, T. W., and Miller C. I.: Evaluating work performance. Am J Clin Nutr 24: 1171–1179 (1971).
8. LaPorte, R. E., et al.: The epidemiology of physical activity in children, college students, middle aged men, menopausal females and monkeys. J Chronic Dis 35: 787–795 (1982).
9. Paffenbarger, R. S., and Hale, R. T.: Exercise and the prevention of coronary heart disease. Prev Med 13: 3–22 (1984).
10. Montoye, H. J., and Taylor, H. L.: Measurement of physical activity in population studies: a review. Hum Biol 56: 195–216 (1984).
11. Edholm, O. C., Fletcher, J. G., Widdowson, E. M., and McCance, R. A.: The energy expenditure and food intake of individual man. Br J Nutr 9: 286–300 (1955).
12. Acheson, K. J., et al.: The measurement of daily energy expenditure—an evaluation of some techniques. Am J Clin Nutr 33: 1155–1164 (1980).
13. Passmore, R., and Durnin, J. V. G. A.: Human energy expenditure. Physiol Rev 35: 801–840 (1955).
14. Durnin, J. V. G. A., and Namyslowski, L.: Individual variations in energy expenditure of standardized activities. J Physiol 143: 573–578 (1958).
15. Bouchard, C.: Reply to letter by Torun. Am J Clin Nutr 38: 815 (1983).
16. LaPorte, R. E., et al.: An objective measure of physical activity for epidemiologic research. Am J Epidemiol 109: 158–168 (1979).
17. Morris, J. N., Everitt, M. G., Pollard, R., and Chave, S. P. W.: Vigorous exercise in leisure-time: protection against coronary heart disease. Lancet 8206: 1207–1210, Dec. 6, 1980.
18. Yasin, S.: Measuring habitual leisure time physical activity by recall record questionnaire. *In* Physical activity and the heart, edited by M. J. Karvonen and A. J. Barry, Charles C Thomas Co., Springfield, IL, 1967, pp. 372–373.
19. Alderson, M., and Yasin, S.: Measuring habitual leisure time activity: a questionnaire method suitable for epidemiologic studies. *In* Physical activity in health and disease, edited by K. Evang and K. Lange-Andersen. Williams and Wilkins Co., Baltimore, 1966, pp. 215–221.

20. Marr, J. W., et al.: Diet, leisure activity and skinfold measurements of sedentary men. Proc Nutr Soc 29: 17a (1970).
21. Bouchard, C., et al.: A method to assess energy expenditure in children and adults. Am J Clin Nutr 37: 461–467 (1983).
22. Paffenbarger, R. S., Wing, A. L., and Hyde, R. T.: Physical activity as an index of heart attack risk in college alumni. Am J Epidemiol 108: 161–175 (1978).
23. Blair, S. N.: How to assess exercise habits and physical fitness. *In* Behavior and health: a handbook of health enhancement and disease prevention; edited by J. D. Matarazzo, et al. John Wiley and Sons, New York, 1984, pp. 424–447.
24. Sallis, J. F., et al.: Physical activity assessment methodology in the Stanford five-cities project. Am J Epidemiol 121: 91–106 (1985).
25. Taylor, H. L., et al.: A questionnaire for the assessment of leisure time physical activities. J Chronic Dis 31: 741–755 (1978).
26. Montoye, H. J.: Estimation of habitual physical activity by questionnaire and interview. Am J Clin Nutr 24: 1113–1118 (1971).
27. Folsom, A. R., et al.: Distribution of leisure time physical activity and its relationship to coronary risk factors in a metropolitan area: the Minnesota Heart Survey. Am J Epidemiol. In press.
28. Shapiro, S., Weinblatt, E., Frank, C. W., and Sager, R. V.: The H. I. P. study of incidence and prognosis of coronary heart disease: preliminary findings on incidence of myocardial infarction and angina. J Chronic Dis 18: 527–558 (1965).
29. Kannel, W. B., and Sorlie, P. S.: Some health benefits of physical activity: the Framingham Study. Arch Intern Med 139: 857–861 (1979).
30. Salonen, J. T., Puska, P., and Tuomilheto, J.: Physical activity and risk of myocardial infarction, cerebral stroke, and death. Am J Epidemiol 115: 526–537 (1982).
31. Grimby, G., et al.: Habitual physical activity: aerobic power and blood lipids. *In* Muscle metabolism during exercise, edited by E. D. Pernow and B. Saltin. Plenum Press, New York, 1979, pp. 469–481.
32. Magnus, K., Matroos, A., and Strackee, J.: Walking, cycling, or gardening, with or without seasonal interruption, in relation to acute coronary events. Am J Epidemiol 110: 724–733 (1979).
33. Baecke, J. A. H., Burema, J., and Frijters, J. E. R.: A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr 36: 932–942 (1982).
34. Morrison, J. F., van Malsen, S., and Noakes, T. D.: Leisure-time physical activity levels, cardiovascular fitness and coronary risk factors in 1015 white Zimbabweans. South Afr Med J 65: 250–256 (1984).
35. Haskell, W. L., et al.: Strenuous physical activity, treadmill exercise test performance and high density lipoprotein cholesterol. Circulation 62 (Supp. IV): 53–61 (1980).
36. Bloom, B.: Current estimates from the National Health Interview, United States, 1981. Vital and Health Statistics, Series 10, No. 141. DHHS Publication No. (PHS) 83–1569. National Center for Health Statistics, Hyattsville, MD, October 1982.
37. LaPorte, R. E., et al.: The spectrum of physical activity, cardiovascular disease and health: an epidemiologic perspective. Am J Epidemiol 120: 507–517 (1984).
38. Baranowski, T., et al.: Validity of children's self-report of aerobic activity. Fam Health Proj Res Q Exerc Sport 55: 309–317 (1984).
39. DeBacker, G., et al.: Physical activity and physical fitness levels of Belgian males aged 40–55 years. Cardiology 67: 110–128 (1981).
40. Klissouras, V.: Heritability of adaptive variation. J Appl Physiol 31: 338–344 (1971).
41. Klissouras, V.: Genetic aspects of physical fitness. J Sports Med 13: 164–170 (1973).
42. Froelicher, V. F., and Oberman, A.: Analysis of epidemiologic studies of physical inactivity as a risk factor for coronary artery disease. Prog Cardiovasc Dis 15: 41–65 (1972).
43. Schoeller, D. A., and van Santen, E.: Measurement of energy expenditure in humans by doubly labeled water method. J Appl Physiol 53: 955–959 (1982).
44. Klein, P. D., et al.: Calorimetric validation of the doubly-labeled water method for estimation of energy expenditure in man. Hum Nutr Clin Nutr 38C: 95–106 (1984).
45. Schoeller, D. A., and Webb, P.: Five day comparison of the doubly labeled water method with respiratory gas exchange. Am J Clin Nutr 40: 153–158 (1984).
46. Schoeller, D. A.: Energy expenditure from doubly labeled water: some fundamental considerations in humans. Am J Clin Nutr 38: 999–1005 (1983).
47. Wade, M. G., and Ellis, M. J.: Measurement of free-range activity in children as modified by social and environmental complexity. Am J Clin Nutr 24: 1457–1460 (1971).
48. Klesges, R. C., Coates, T. J., and Molderhauer, L. M.: The FATS: an observational system for assessing physical activity in children and associated parent behavior. Behav Assessment 6: 333–345 (1984).
49. Bullen, B. A., Reed, R. B., and Mayer, J.: Physical activity of obese and non-obese adolescent girls appraised by motion picture sampling. Am J Clin Nutr 14: 211–223 (1964).
50. Booyens, J., and Hervey, G. R.: The pulse rate as a means of measuring metabolic rate in man. Can J Biochem Physiol 38: 1301–1309 (1960).
51. Taylor, C. B., et al.: A new system for long-term recording and processing of heart rate and physical activity in outpatients. Comput Biomed Res 15: 7–17 (1982).
52. Bradfield, R. B.: A technique for determination of usual daily expenditure in the field. Am J Clin Nutr 24: 1148–1154 (1971).
53. Warnold, T., and Lenner, R. A.: Evaluation of the heart rate method to determine the daily energy expenditure in disease. A study of juvenile diabetics. Am J Clin Nutr 30: 304–315 (1977).
54. Vokac, Z., Bell, H., Bautz-Holter, H., and Rodahl, K.: Oxygen uptake/heart rate relationship in leg and arm exercise, sitting and standing. J Appl Physiol 39: 54–59 (1975).
55. McCloskey, D., and Streatfield, K.: Muscular reflex stimuli to the cardiovascular system during isometric contraction of muscle groups of different mass. J Physiol 230: 431–441 (1975).
56. Sengupta, A., Saka, D., Muklonadhyay, S., and Gosivamn, P.: Relationship between pulse rate and energy expenditure during graded work at different temperatures. Ergonomics 22: 1207–1215 (1979).
57. Kappagoda, C., Linden, R., and Newell, J.: Effects of the Canadian Air Force Training Program on submaximal exercise test. Q J Exp Physiol 64: 185–204 (1979).
58. LeBlanc, J. A.: Use of heart rate as an index of work output. J Appl Physiol 10: 275–280 (1967).

59. Bateman, S. et al.: Heart rate of training captains engaged in different activities. Aerospace Med 41: 425–429 (1970).
60. Christensen, C. C., et al.: A critical evaluation of energy expenditure estimates based on individual O_2 consumption/heart rate curves and average daily heart rate. Am J Clin Nutr 37: 468–472 (1983).
61. Washburn, R. A., and Montoye, H. J.: Validity of heart rate as a measure of mean daily energy expenditure [abstract]. Med Sci Sports Exerc 16: 196–197 (1984).
62. Goldsmith, R., and Hale, T.: Relationship between habitual physical activity and physical fitness. Am J Clin Nutr 24: 1489–1493 (1971).
63. Garrow, J. S.: Energy balance and obesity in man. Elsevier/North Holland Publishing Biomedical Press, Amsterdam, 1974, p. 63.
64. Sander, L. W., and Julia, H. L.: Continuous interactional monitoring in the neonate. Psychosom Med 28: 822–835 (1966).
65. Sander, L. W., Julia, H. L., Stechler, G., and Burns, P.: Continuous 24-hour interactional monitoring in infants reared in two caretaking environments. Psychosom Med 34: 270–283 (1972).
66. Rose, H. E., and Mayer, J.: Activity, caloric intake, fat storage, and energy balance of infants. Pediatrics 41: 18–29 (1968).
67. Bloom, W. L., and Eidex, M. F.: Inactivity as a major factor in adult obesity. Metabolism 8: 679–684 (1967).
68. Stunkard, A.: A method of studying physical activity in man. Am J Clin Nutr 8: 595–601 (1960).
69. Gayle, R. H., Montoye, H. J., and Philpot, J.: Accuracy of pedometers for measuring distance walked. Res Q 48: 632–636 (1977).
70. Washburn, R. A., Chin, M. K., and Montoye, H. J.: Accuracy of pedometer in walking and running. Res Q Exerc Sports 51: 695–702 (1980).
71. Saunders, J., Inman, V., and Eberhart, T.: The major determinants in normal and pathological gait. J Bone Joint Surg 35A: 543–558 (1953).
72. Marsden, J. P., and Montgomery, S. R.: A general survey of the walking habits of individuals. Ergonomics 15: 429–441 (1972).
73. Black-Sandler, R., et al.: Determinants of bone mass in the menopause. Prev Med 11: 269–280 (1982).
74. Brugger, W., and Milner, M.: Computer-aided tracking of body movement using A. C. C. D. usage sensor. Med Biol Eng Comput 16: 207–210 (1978).
75. Morris, J. R. W.: Accelerometry—a technique for measurement of human body movements. J Biomech 6: 729–736 (1973).
76. Schulman, J. L., Stevens, T. M., and Kupst, M. J.: The biomotometer: a new device for the measurement and remediation of hyperactivity. Child Dev 48: 1152–1154 (1977).
77. Wong, T. C., Webster, J. G., Montoye, H. J., and Washburn, R. A.: Portable accelerometer device for measuring human energy expenditure. IEEE Trans Biomed Eng BME-28: 467–471 (1981).
78. Foster, F. G., McPartland, R. J., and Kupfer, D. J.: Motion sensors in medicine: a report on reliability and validity. J Inter-Amer Med 3: 4–8 (1978).
79. Cauley, J. A., et al.: Physical activity and HDL–C subfractions in post-menopausal women. Paper presented at the American College of Sports Medicine, San Diego, May 1984 [abstract]. Med Sci Sports Exerc 16: 176 (1984).
80. Montoye, H. J., et al.: Estimation of energy expenditure by a portable accelerometer. Med Sci Sports Exerc 15: 403–407 (1983).
81. Sopko, G., Jacobs, D. R., and Taylor, H.: Dietary measures of physical activity. Am J Epidemiol 120: 900–911 (1984).
82. Beaton, G. H., et al.: Source of variance in 24 hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins, and minerals. Am J Clin Nutr 37: 986–995 (1983).
83. Buskirk, E. R., Harris, D., Mendez, J., and Skinner, J. Comparison of two assessments of physical activity and a survey method for caloric intake. Am J Clin Nutr 24: 1119–1125 (1971).
84. Caspersen, C. J., Powell, K. E., and Christenson, G. M.: Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 100: 126–131, March-April 1985.
85. Haskell, W. L., Montoye, H. J., and Orenstein, D.: Physical activity and exercise to achieve the health-related components of physical fitness. Public Health Rep 100: 202–212, March-April, 1985.
86. Siscovick, D. S., LaPorte, R. E., and Newman, J. M.: The disease-specific benefits and risks of physical activity and exercise. Public Heath Rep 100: 180–188, March-April 1985.
87. LaPorte, R. E., et al.: The assessment of physical activity in older women: analysis of the interrelationship and reliability of activity monitoring, activity surveys and caloric intake. J Gerontol 38: 394–397 (1983).
88. Cauley, J. A., LaPorte, R. E., and Black-Sandler, R.: Measurement of physical activity in older populations. Paper presented at Society for Epidemiologic Research, Houston, 1984 [abstract]. Am J Epidemiol 120: 471 (1984).