TO INVESTIGATE THE ROLE of physical activity in maintaining health, the traditional argument has been that the number of outcome events is limiting, and therefore large numbers of individuals must be enrolled and costs per assessment must be small. Based on both this rationale and the fact that physical activity has been a secondary variable of interest in early studies, many large cohort studies have been performed using self-reported physical activity recalls. These studies have provided strong and convincing evidence that physical activity can protect health by lowering the risk for coronary heart disease, stroke, hypertension, Type II diabetes, metabolic syndrome, depression, and colon and breast cancers (12). Because these overarching questions have been largely answered, investigators have turned their attention to important questions regarding dose-response relationships; health protection in smaller, more homogeneous cohorts; and the effectiveness of interventions to increase physical activity behaviors. These questions are likely to require more precision and accuracy in measurement than can be provided by the traditional survey methods.

Validation studies using objective measures of physical activity or energy expenditure indicate that the accuracy and precision of survey techniques are limited. A recent review of physical activity surveys as estimates of energy expenditure concluded that none of the 23 questionnaires evaluated had both acceptable correlations and mean differences compared to doubly labeled water (DLW) measures of energy expenditure at the group level, as well as reasonable individual estimates of the total energy expended in physical activity (7). Intensity of activity plays an important role in the accuracy of physical activity recalls, with reasonable accuracy and precision for vigorous physical activity, but not for less intense activities (1, 4, 5). Convincing evidence to date for the inadequacy of our current questionnaires comes from examinations of physical activity in the 2003–2004 and more recently, the 2005–2006 NHANES study in which the physical activity of U.S. adults was measured by accelerometer (10, 11). The proportion of adults self-reporting enough physical activity to meet current recommendations was from 6- to 10-fold higher than that measured by accelerometer. This substantial difference has also been found in the applied setting, where, for example, obese patients undergoing bariatric surgery self-report five-fold increases in their moderate-vigorous activity postsurgery, while accelerometer-determined measures indicate a net decrease in their physical activity (2). Certainly the cut point selected to define moderate or greater intensity activity can affect these proportions, but although the absolute values may change with changes in these cut points, having the raw accelerometer output allows one to apply different scoring methods to this objective data. Our own data in which we directly compared various objective and survey methods of measuring physical activity demonstrated that all three objective measures correlated better with DLW-derived PAEE and had lower error than any of the three surveys (Fig. 1). Clearly whether for surveillance or detection of change in activity in a particular setting, our current self-report methodologies lack precision and accuracy and thus may lead to faulty conclusions and overestimated recommendations regarding the dose of physical activity needed to maintain health.

A criterion method for the validation of physical activity questionnaires is DLW, which is an accurate and reasonably precise measure of total energy expenditure (8). We have reported a very strong relationship of activity energy expenditure (AEE) derived from DLW with decreased mortality in only 302 subjects (6), with the self-reported physical activity of these same subjects not corresponding well with their measured AEE (6). DLW, however, has major limitations. One basic argument is that it only measures the energy cost of physical activity and not the behavior itself. Thus DLW does not provide relevant information regarding frequency, intensity, duration, pattern, and type of activity. Some of these components may be important predictors of specific health outcomes and knowledge of these will facilitate better physical activity recommendations. DLW is also quite expensive and requires specific expertise for its use. The cost limitation, however, is mediated in many applications, because the precision allows investigators to obtain statistically significant results with a smaller sample. Power calculations based on the precision of self-report instruments indicated that a sample size of 1,000–1,200 individuals would have been required for our study of AEE and mortality.

During the past decade there has been an explosive increase in the number and variety of commercially available objective physical activity monitors. There are dozens of accelerometers on the market, which provide information on the frequency, intensity, and duration of activity as well as the time pattern. Moreover, recent work suggests that the raw data may provide sufficient information to accurately identify the type of activity (9). Drawbacks to accelerometers include cost, requirements for direct interaction between researchers and subjects, or at least a means to mail the monitors, and an inability to capture all forms of physical activities. Thus accelerometers may not be useful in large studies, but again, because they are objective instruments, accurate and reasonably precise, many of the current questions about the health benefits of physical activity may not require large cohorts. Although accelerometers cannot capture all activities (water-based activity, cycling, upper-body...
univariate regression. 

or resistance exercise) and are limited to the assessment of current activity, in combination with a simple physical activity log they have great utility, particularly in smaller intervention studies. Declines in cost and an increased awareness of their capabilities will make them more appealing in larger studies.

If even modest cost is an impediment to using objective monitors in studies of physical activity, the cost and quality of pedometers have undergone revolutionary changes, with costs having dropped below $20 with bulk purchases. The least expensive models are limited in much the same way as DLW in that they count only total steps and thus do not provide information on intensity, frequency, duration, or pattern through the day, but at least in an elderly cohort, they correlate well with total energy expenditure by DLW and vastly outperform survey techniques in this regard (Fig. 1; Ref. 3). For about $30 each, there are pedometers available that include a time stamp that obviates many of the above limitations. They provide pattern through the day and step frequency and thus field testing and validation are needed.

While there are still needs for traditional survey techniques such as the assessment of historical levels of activity among adults, or the assessment of resistance exercise, swimming, biking or activities in rough terrain, the increase in accuracy and precision available with the use of monitors reduces the value of traditional survey methods for most studies (3, 10). Coupled with dramatic price reductions, even the cost argument in support of traditional survey methods is weakening. It is our belief a tipping point has been reached and most investigations should be using monitors because of their superior accuracy and precision.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

REFERENCES