The assessment of physical activity in individuals and populations: Why try to be more precise about how physical activity is assessed?

Article *in* International Journal of Obesity · September 1998

Source: PubMed

Kirsten Rennie
University of Cambridge

International Journal of Obesity (1998) 22 Suppl 2, S30–S38

http://www.stockton-press.co.uk/ijo

The assessment of physical activity in individuals and populations: Why try to be more precise about how physical activity is assessed?

NJ Wareham and KL Rennie

Department of Community Medicine, University of Cambridge, Cambridge, CB2 2SR, UK

Simple epidemiological measures of physical activity have proved sufficient to demonstrate associations with many chronic disease outcomes, but they have infrequently separated physical activity into its different dimensions, nor have they allowed estimation of dose-response effects. Generating greater clarity about the nature of the exposure-disease relationship, is an important step in the development of an appropriate public health intervention. This clarity can only be achieved with reliable and valid measurement instruments, which objectively and quantitatively assess the dimension of physical activity that is of interest for a particular health outcome. Objective techniques, such as heart rate monitoring, which have been directly compared to gold standard assessment methods, may be of use in medium-sized epidemiological studies and as a validation tool for questionnaires to be used in larger studies. The combination of methods with uncorrelated error, would result in an improved estimation of the true exposure and is an important area for research. Improved assessment would be of use in aetiological studies, in tracking trends in physical activity within populations, making objective comparisons between populations and in monitoring the effect of interventions.

Keywords: physical activity; energy expenditure; fitness; exercise; epidemiology

Introduction

Much of the evidence regarding the role of reduced physical activity as a risk factor for human disease, comes from epidemiological investigations. Although a large number of studies have demonstrated links to all-cause mortality and to a wide range of specific disease endpoints, including coronary heart disease, osteoporosis, diabetes and some cancers,[1] the methods used to assess physical activity have been relatively crude. The goal of any epidemiological study is to define the strength of the association between the exposure (in this case physical activity) and the outcome, and to judge the extent to which that association is likely to be causal and not simply a function of chance, bias or confounding. The selection of the method to assess the exposure in a population is critical to these studies and is a delicate balance between feasibility and accuracy. Traditionally epidemiologists have tended to put more emphasis on feasibility. They argue that provided there is no bias in the estimation of the exposure, the observed relationship with a given outcome will tend to be an under-estimation of the true association and that this will be statistically significant if the sample size is large enough.[2] Although this approach is probably sufficient for simple binary exposures, it is inadequate for the study of complex exposures like physical activity. Using the example of the association between physical activity and diabetes, we will provide five reasons to explain why epidemiologists need to develop more accurate and objective methods to assess physical activity in population-based studies. Overall there has been a wide separation between the methods used in epidemiological studies and those more accurate techniques that are applicable to smaller studies. This review describes the range of methods that are available and discusses possible approaches to narrowing the gap between feasibility and accuracy.

One particular problem in physical activity epidemiology, is that different people use common terms to mean different things. For example, it is common for authors to imply that they have measured energy expenditure, (EE) when what they have really measured is self-reported participation in sports and recreational activity. Therefore, in the interests of clarity, we will use terms as defined by Caspersen and colleagues[1,3] (Figure 1)

Correspondence: Dr NJ Wareham, Department of Community Medicine, University of Cambridge, Institute of Public Health, Robinson Way, Cambridge, CB2 2SR, UK.

Subjective assessment of physical activity

A wide variety of different questionnaire-based methods have been used to describe physical activity in

Physical activity, is any bodily movement produced by the contraction of skeletal muscles resulting in caloric expenditure.
Exercise, is a sub-category of physical activity and is activity which is planned, structured and repetitive.
Physical fitness, by contrast, is a set of outcomes or traits that relate to the ability to perform physical activity.
Cardio-respiratory fitness, is a health-related component of physical fitness that relates to the ability of the circulatory and respiratory systems to supply oxygen during sustained physical activity.

Figure 1 Definitions of terms used in physical activity epidemiology.

epidemiological studies. The questionnaires differ in the way that they are administered, the target population in which they can be used, the time frame over which activity is assessed, the type of activity which is measured and the scale to which this exposure is reduced.[4] Some researchers attempt to reduce one of the most complex exposures in epidemiology to a single question. Others try to assess the exposure in a more complete manner and then reduce it to a quantitative index. It is not the purpose of this paper to describe the whole range of questionnaires, as these have been reviewed elsewhere and reproduced in recent publications.[5] However, in this context it is important to give an overall picture of what inferences can be drawn from epidemiological associations that are assessed using such simple subjective measures.

What underlying exposure do questionnaires assess?

One criticism of questionnaires is that they tend to focus on only one aspect of everyday activity and few have been developed to assess all the major areas of physical activity (exercise, sports and recreations, work, and non-sporting, non-occupational activity).[6] Overall, questionnaires are most effective at measuring easily recalled programmed activities. Participation in sports and programmed exercise such as weight training, swimming, jogging or aerobic dance is possible to assess by questionnaire because the individual makes a purposeful decision to undertake the activity, which is discrete and time limited, all of which are factors that aid recall. The energy expended in these activities can be quantified by ascribing an energy cost to the particular pursuit (from published compendia),[7] identifying the time taken for the activity and then multiplying this by the frequency with which the activity was undertaken during the time period of interest. This approach is common to many questionnaires. However, the focus on sports and recreations is driven by the ease with which this sub-category of physical activity can be assessed rather than by its importance.

Another problem with physical activity questionnaires is that it is sometimes unclear which of the underlying dimensions of this complex exposure is being assessed. As EE during sports and recreations is only a small component of total EE (TEE), it is unlikely that self-reported participation in this limited range of activities could be an accurate assessment of the totality. Because sports are often vigorous conditioning activities, it is more likely that assessing participation in sports provides an indication of cardio-respiratory fitness. Generating clarity about what underlying physiological exposure is being assessed, becomes important when epidemiological information is translated into recommendations for public health intervention. If participation in sports and recreations predicts a particular health outcome, can we be sure that we need to recommend increased vigorous activity or is it possible that simply increasing overall EE might produce the same effect? This has important public health implications, as it is a very different proposition to recommend increased participation in high intensity activities compared to increasing TEE, which can be achieved with a more frequent lower intensity activity, such as walking.

The design of validation studies for physical activity questionnaires

Another broad area of criticism of physical activity questionnaires relates to how their validity is demonstrated. Ideally the comparison or gold standard method against which a questionnaire is compared should assess the true underlying exposure in the same frame of reference as the questionnaire. Thus if the objective of a questionnaire is to estimate total energy expenditure in the past week, then the gold standard method should be an objective measurement of TEE over this time. Much rests on the selection of the gold standard method. Comparison between different questionnaires provides information about whether they give the same answer, but little information about whether either of them is valid.[8] Neither is it sufficient to compare questionnaires against other subjective measurement instruments, which are likely to have correlated error. This may be particularly true of comparisons of questionnaires with self-completed physical activity diaries.[9,10] If it is easier to recall participation in vigorous activities, then the same recall bias will be apparent in a questionnaire as in a diary. Therefore, as these errors are correlated, both methods may give similar results. The choice of a gold standard method for validating questionnaires whose frame of reference is long is particularly difficult. In these situations, the true underlying exposure of interest, (for example, habitual energy expenditure in the past year) is a latent variable, as it is unmeasurable even with the best gold standard method. Therefore, the most suitable approach is to estimate the latent variable using repeated measurements with a gold standard method.[11] Few validation studies have taken this approach, principally due to the cost. In the Survey of Activity Fitness and Exercise (SAFE) study,[6] ten physical activity questionnaires were compared with two-day physical activity diaries, collected monthly over 14 months and an assessment using an accelerometer. At each visit, the subjects completed a Four-Week History questionnaire modified from the Minnesota Leisure Time Physical Activity questionnaire[12] and other measures, including body fatness by underwater weighing, skinfold thickness at five sites, VO_{2max} and forced expiratory volume, were taken. Although this was a relatively large and intricate study, the choice of gold standard methods can be criticised as either likely to have correlated error or being too distant from the true exposure of interest. The markers of obesity, fitness and respiratory function might be associated with physical activity, but they are not direct measurements of it, so cannot be considered as true gold standard methods.

The selection of populations for validation studies

A second key requisite of an appropriate validation study is that the subjects who are recruited should be, as far as possible, representative of the population to whom the questionnaire will be administered. In the SAFE Study[6] the 78 study participants (50 women and 28 men) were mostly highly-educated Caucasian University employees (71% had college or graduate degrees). They were atypical by comparison to the general population of the USA in that 95% were non-smokers, 85% were employed full-time in sedentary occupations and 45% rated themselves as vigorously active more than twice a week. The study population is likely to be enriched with well-educated individuals who were aware of the health benefits of activity and who had a higher level of participation in vigorous leisure-time activity than the general population, and for whom this represented a greater proportion of their TEE by virtue of their sedentary occupations. If leisure-time activity is the easiest component of TEE to assess, then the selection of a population in whom this was the major component of overall activity, may result in an over-estimation of the validity of the questionnaire as a measure of total activity. The major conclusion from this and other studies of the numerous questionnaires that are used in field studies is that, although they demonstrate significant associations with a wide variety of disease end-points, there is major uncertainty over what exactly is being measured. Although the questionnaires correlate fairly strongly with other assessments of vigorous or moderately intense physical activity, the measurement of light to moderate intensity activity is generally less accurate.

Objective assessment of physical activity

There are a number of alternatives to these subjective measurement instruments, but these are rarely used in epidemiological studies.

Direct physiological assessment

The doubly-labelled water method is a non-invasive method for measuring EE in free-living people over a period of 10–20 d. It involves the administration of isotopes of hydrogen and oxygen and the determination of the washout kinetics of both isotopes as their concentrations fall back down to the pre-administration level.[13] Although the technique is feasible and non-invasive, it has not been widely used because of the cost and scarcity of the isotopes ^{2}H and ^{18}O. A recent meta-analysis of the literature reported a total of only 1614 published and unpublished doubly-labelled measurements, over half of which were either repeat estimates or had been conducted in special physiological states such as pregnancy, athletic or military training.[14] However, the technique remains the gold standard method for estimating TEE in free-living individuals.

Whole body calorimetry

An alternative gold standard method is whole body calorimetery, in which EE can be measured by collecting all the expired gases from a subject who lives within a sealed room.[15] Although this technique is very precise, it does not assess the individual in a free-living state and therefore has limited application in epidemiological studies.

Movement sensors

A wide range of movement sensors have been used in physical activities studies to objectively estimate movement. These range from simple pedometers[16] (which can be used to count steps), to more sophisticated instruments, which predict basal metabolic rate (BMR) from an equation based on age, gender and body mass index (BMI) and then add to that an estimate of the energy cost of activity calculated from movement detected by a piezometer, which measures vertical acceleration.[17–19] The latter technique has been shown to be valid in laboratory comparisons,[17] but has been less successful as a single technique for measuring energy expenditure in free-living adults.[20] However, because of the principle that precision can be improved by assessing exposure using combined methods provided they have uncorrelated error, it is likely that movement sensing could have a role when used in conjunction with other methods, although early attempts at this, using less sophisticated equipment, were not very successful.[21,22] Recently, a new triaxial movement sensor has become available, but as yet, full evaluation has not been completed.[23,24]

Heart rate monitoring (HRM)

It has long been known that there is a relationship between EE and heart rate, and many attempts have been made to use recorded heart rate as a means of estimating physical activity. Some methods have used average pulse rate,[25] while others have used linear predictions.[26] Initially, these linear predictions were made using the same parameters for all subjects.[26] However, the relationship between EE and heart rate is dependent upon factors such as body weight and fitness. The most accurate predictions have been obtained using individual calibrations, where the relationship between measured EE and heart rate is defined in each individual. Most of the debate regarding this technique has concentrated on the difficulty of predicting EE when the heart rate is low.[27] A variety of different methods have been devised to circumvent this problem, by using different prediction equations above and below the point of non-linearity. Early dissatisfaction with the technique of HRM, may have been due to equipment problems and the use of single line calibrations.[26–28] As Davey-Smith and Morris[29] have pointed out, however, this technique has considerable potential for use in epidemiology, provided the technical problems can be overcome.

HRM with individual calibration: The flex heart rate (HRFlex) method

This method relies on the fact that there is a linear relationship between heart rate and oxygen consumption above a definable critical level.[30] Below this level, which was termed HRFlex, the relationship is more variable. Therefore, in order to predict EE from heart rate, the linear prediction is used above HRFlex and the average of a series of recordings of resting EE (REE) is used below it. The HRFlex point is empirically defined as the average of the lowest heart rate during exercise and the highest at rest. The method therefore relies on an individual calibration in which REE and heart rate are recorded. The subject then exercises and the relationship between EE and heart rate is monitored (see Figure 2). Four individual parameters are derived from the calibration. These are the slope and intercept of the linear part of the EE relationship, the HRFlex and the average REE. Using these four parameters, EE can be computed from heart rate. The advent of lightweight heart rate monitors, which are worn over the chest, together with recording devices worn like a wrist watch, have allowed heart rates to be monitored and recorded minute-by-minute for extended periods of time without major inconvenience to the individual.[31] This technique, therefore, provides a means of objectively measuring EE. Not only can TEE be computed, but the data can be analysed to look at the pattern of EE. Using indirect calorimetry and the doubly labelled water techniques

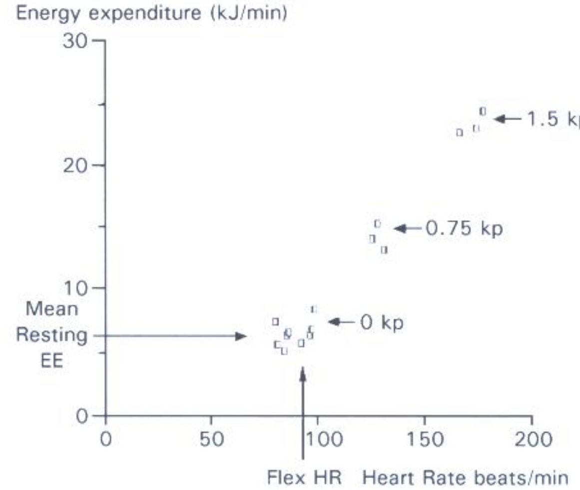

Figure 2 Typical relationship between heart rate (HR) and energy expenditure (EE) as measured during individual calibration.

to accurately measure individual EE, it has been shown that the HRM technique is accurate, relatively cheap and non-invasive, and therefore potentially suitable for epidemiological studies.[30,32–34] Spurr *et al*[30] compared this method with total daily EE measured by whole-body indirect calorimetry in 22 subjects and reported a strong correlation (0.92). Ceesay *et al*[34] in a similar study, compared HRM against whole body calorimetry and concluded that the HRM method yielded a mean non-significant underestimate of total energy expenditure. The results of this study were combined with the results of the Spurr *et al* paper,[30] to give a regression coefficient of 0.96 and an intercept of 0.39 mJ. The mean error of estimating TEE by this technique was only 0.6%. Figure 3 shows the combined results of the studies by Spurr *et al*[30] Ceesay *et al*[34] with the estimate of EE by HRM plotted against that derived from indirect calorimetry or doubly-labelled water.

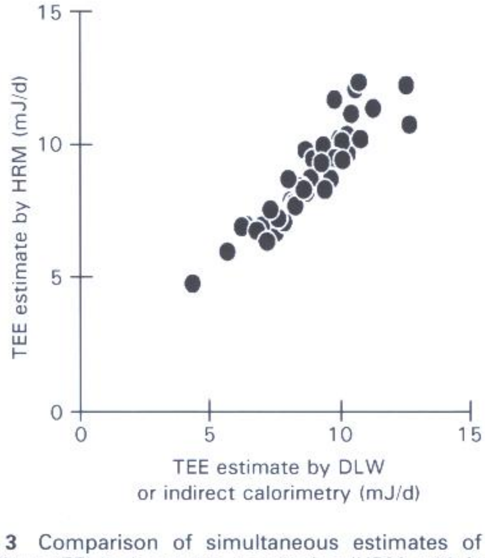

Figure 3 Comparison of simultaneous estimates of energy expenditure (EE) by heart rate monitoring (HRM, with individual calibration) and doubly-labelled water (DLW) or whole body calorimetry. TEE = total EE.

As it is a non-invasive objective method, with a known relationship to an appropriate gold standard, the HRFlex HRM method, has considerable potential as a measurement tool for small to medium-sized epidemiological studies.[31] It can also act as a reference method for questionnaires to be used in larger studies, by being repeated intermittently during the frame of reference, to estimate the latent variable habitual EE. An important area of development of the HRFlex method is the combination of a simplified and technically improved heart rate monitor with a movement sensor. This would allow heart rate acceleration due to anxiety and stress to be more easily differentiated from that related to activity. The combination of two independent and objective methods for measuring physical activity would be a significant advance especially as they are likely to have uncorrelated error.

Why try to be more precise about how physical activity is assessed?

Traditionally, epidemiologists have relied on the assumption that the use of simple exposure measures leads to an underestimation of the true exposure-disease relationship provided the error in the measurement is non-differential. In the case of physical activity, this would imply that the underlying relationships with disease endpoints, are stronger than those that are observed with simple exposure measures. Although this approach is sufficient to demonstrate the overall importance of physical activity as a factor relating to human health, it is inadequate as a basis for generating greater clarity about the exposure-disease relationship. Five reasons which justify research aimed at improving the methods for assessing physical activity are shown in Figure 4.

- To specify which aspect of physical activity is important for a particular health outcome.
- To estimate more accurately the effect size.
- To make cross-cultural comparisons.
- To monitor temporal trends in physical activity.
- To monitor the effect of interventions.

Figure 4 Reasons for seeking to improve the methods for assessing physical activity in epidemiological studies.

1. Specifying which dimension of physical activity is of most importance for a particular health outcome

The translation of epidemiological evidence into public health action, demands considerable clarity about the nature of the disease-exposure relationship. Unlike some adverse exposures (for example, cigarette smoking), physical activity is a complex multi-dimensional exposure and therefore, clear evidence is required to indicate which aspect of the exposure is important for a particular health effect and how much activity is likely to have an effect. The recent report on Physical Activity and Health from the Surgeon General in the United States[1] concluded that further research was required to 'delineate the most important features or combinations of features of physical activity . . . that confer specific health benefits'. To take the example of non-insulin-dependent diabetes (NIDDM), there are seven prospective cohort studies, all of which conclude that physical activity is protective (Table 1).[35–41] The observed associations between physical activity and diabetes are consistent and are not confounded by other known risk factors for diabetes, such as age and obesity. However, these studies use questionnaire-based methods for assessing physical activity, and none of them measure

Table 1 Prospective studies of the association between physical activity and the incidence of non-insulin dependent diabetes mellitus (NIDDM)

Study and reference	*Subjects*	*Assessment of physical activity*
US Nurses' Health Study[35]	87 253 women aged 34–59 y studied for eight years. 1303 new cases of NIDDM	Single question on frequency of sweat-inducing vigorous activity
University of Pennsylvania Alumni Health Study[36]	5990 men aged 39–68 y studied for 14 years. 202 new cases of NIDDM	Paffenbarger questionnaire
Physicians' Health Study[37]	21 271 men aged 40–84 y studied for five years. 285 new cases of NIDDM	Single question on frequency of sweat-inducing vigorous activity
British Regional Heart Study[38]	7735 men aged 40–59 y studied for 12.8 years. 194 new cases of NIDDM	Physical activity score in six levels
Honolulu Heart Program[39]	6815 Japanese American men aged 45–68 y. 391 new cases of NIDDM	h/d spent on each of five different activity levels
Kuopio Ischemic Heart Disease Risk Factor Study[40]	897 men aged 42–60 y studied for four years. 46 new cases of NIDDM	Variant of the Minnesota leisure time physical activity questionnaire
Tampere Study[41]	891 men and 973 women aged 35–63 y studied for 10 years. 118 new cases of NIDDM	23 questions on conditioning exercise, sports, leisure time and household chores

occupational, as well as recreational, EE–design features which make interpretation of their results difficult. It also potentially hinders the process of translating their results into public health intervention. For example, the assessment of exposure in the Nurse's Health Study[35] and US Male Physicians Study,[37] is on the reported frequency of vigorous exercise per week, and is based on the response to the question 'at least once a week do you engage in any regular activity similar to brisk walking, jogging, bicycling etc, long enough to work up a sweat?'. This question is said to be have been validated as a measure of physical activity. However, analysis of these validation studies suggests that some doubt exists about what exactly is being measured by such a question. In one of the cited validations, Washburn *et al*[42] compared the response to this question with the score derived from the Harvard Alumni Activity Survey Questionnaire, the BMI and the high density lipoprotein (HDL) cholesterol in 732 randomly selected adults. The correlation with the Harvard Questionnaire was 0.24. In a second 'validation', the results of the sweat-induced activity question were compared with results from a maximal oxygen uptake fitness test.[43] A correlation of 0.46 was reported between the reported number of days per week with exercise-induced sweating and VO_{2max} per kg in 78 volunteers. The lack of clarity about which dimension of physical activity the questionnaire is assessing, together with the selection of possibly unsuitable validation instruments, makes it difficult to interpret the results of studies using this questionnaire. Similar criticisms can be made of the other studies and thus, although there is apparently strong evidence that physical activity is in some way causally associated with diabetes, there is uncertainty about which dimension of this complex exposure is being assessed.

2. Estimating the true effect size

A second reason for wanting greater precision in the estimation of the different dimensions of physical activity, is to produce information about the magnitude and nature of the exposure–disease relationships. The use of non-quantitative measures of physical activity does not allow dose–response relationships to be examined, nor does it allow detailed examination of relationships for linearity or thresholds. Perhaps even more serious, is the use of semi-quantitative measures of physical activity, from which a dose-response is computed. If the validity of the original measure is questionable or if its focus is only on one aspect of the exposure, then the dose-response relationship may be of limited value. An example of this in relation to diabetes is the physical activity index used in the University of Pennsylvania Alumni Study which is based on the sum of the energy costs of self-reported walking, stair climbing and recreational activity[36] assessed using the Paffenbarger *et al* questionnaire.[44] In the analysis of this study (see Table 1), a 500 kcal/week increase in this index, was associated with a relative risk reduction of 6% in a model that also adjusted for BMI, history of hypertension and parental history of diabetes. The authors inferred from these observations, that physical activity was associated with a reduced risk of NIDDM. The presentation of EE data as kcal or kJ per unit time, without any adjustment for weight, does not take account of differences in body size, which have a major impact on EE.[7]

Therefore, although the data presented are adjusted for BMI, residual confounding by weight could still have occurred. The presentation of the expected reduction of risk for each 500 kcal/week increase in the physical activity index is shown by imposing linearity onto semi-quantitative data. The uncertainties in this quantitative conclusion can create problems if it is used to inform public health programmes or to estimate the likely benefit of interventions.

3. Making cross-cultural comparisons

As many physical activity questionnaires are relevant only to the particular populations for which they were designed, comparisons of physical activity patterns between countries or between different social or cultural groups within a country is difficult. Therefore, it is difficult to undertake formal ecological studies, in which international variation in levels of physical activity are compared to observed differences in disease rates between countries. This form of data would be of use in disease such as diabetes, in which there is marked geographical variation,[45] with low prevalence in traditional rural areas of previously undeveloped countries and high rates in populations in whom traditional lifestyles have been replaced by westernised dietary and physical activity patterns.[46] The cultural specificity of many subjective physical activity instruments is illustrated by questionnaires like that used in the British Regional Heart Study[38] (see Table 1) which focuses on specific pursuits such as sailing, golf and tennis. This restricted focus limits the use of this type of questionnaire, making it unacceptable for comparisons between socially distinct sub-groups within a country and for between-country studies. This criticism does not imply that this questionnaire was inappropriate for the original study for which it was developed, merely that comparison between populations should be made with caution. The single item physical activity question used in the Nurses' and Physicians' Health Studies[35,37] may be successful in predicting outcomes within these studies, but at least part of its utility must stem from the particular characteristics of the cohorts for which it was developed. These populations are socially homogeneous and occupationally defined and, therefore, the major physical activity differences between individuals will be in leisure time activity. It is unclear whether this questionnaire would be applicable to studies which recruit volunteers from diverse social groups or occupations. Overall, comparability between studies would be assisted by the development of objective measures which can be used equally well in populations of different age, social class, occupation and cultural background.

4. Monitoring temporal trends within populations
Such objective measures would also be of use in demonstrating temporal trends in physical activity within populations. The absence of reliable population-level physical activity data is highlighted in comparison with the availability of data on energy intake and food and nutrient consumption. In a recent article discussing the causes for the rapid rise in the prevalence of obesity in the UK, Prentice and Jebb[47] illustrated their argument with reasonable evidence on temporal trends in energy intake and in the proportion of energy originating from dietary fat. However, the inferences about trends in EE were drawn from changes in television watching and car ownership. More direct evidence on trends in physical activity exists in the US[48] and will become available in the UK with the successive publications of the Health Survey for England.[49] However, even these data may fail to demonstrate changes in the true exposures of interest. A clear distinction should be made between population-level data on self-reported behaviour and data on distributions of fitness and EE which are not currently available. Theoretically, an increase in the prevalence of a self-reported behaviour, such as leisure time physical activity, may be balanced by an opposite trend in work-related physical activity, with little effect on overall EE. Only by separating temporal trends in behaviours and underlying physiological exposures will greater clarity be generated.

5. Measuring the effect of interventions
Finally, more objective measures of physical activity would be of use in monitoring the effect of interventions. Although physical activity intervention studies may demonstrate an effect on a particular outcome, monitoring the means by which that effect was achieved is an important goal. As it is difficult to blind participants in a behavioural intervention study to the hypothesis that is being tested, instruments that rely on self-reported behaviour will always be subject to bias. The limited availability of objective measurement instruments has meant that those studies which have attempted to measure change objectively have concentrated on areas for which better techniques exist, such as cardio-respiratory fitness.[50] Demonstrating changes in energy turnover has not been a priority as this is difficult to study. As the goal of recent public health advice is to increase participation in low to moderate intensity activity, which will not result in large changes in fitness, but will increase TEE,[51,52] the development of more objective measures of this exposure is critical to the ability to evaluate the success of these programs.

Conclusions

In a review of the methods employed to assess physical activity, Caspersen[53] identified several key problems which have characterised studies of the link between physical activity and health. These included the diverse definitions of physical activity and exercise, the absence of valid assessment instruments that can be used across studies and the failure to develop instruments that reflect the different health-related components of physical activity. Although progress has been made in refining the measurement of physical activity in epidemiological studies, these problems persist. The use of simple physical activity instruments may have been sufficient to demonstrate the overall importance of this factor in chronic disease causation, but future studies will need to give much greater attention to the precision with which this exposure is measured. The use of single or combined objective measures, which have a known relationship to a physiologically-defined exposure of interest, is a promising area of research.

Acknowledgements
NJW is an MRC clinician scientist fellow.

References
1 US Department of Health and Human Services. *Physical activity and health: A Report of the Surgeon General.* US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion: Atlanta, 1996.
2 Armstrong BK, White E, Saracci R. *Principles of exposure measurement in epidemiology.* Oxford University Press: Oxford, 1994.
3 Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985; **100**: 126–131.
4 Washburn RA, Montoye HJ. The assessment of physical activity by questionnaire. *Am J Epidemiol* 1986; **123**: 563–576.
5 Kriska AM, Caspersen CJ. A collection of Physical Activity Questionnaires for Health-Related Research. *Med Sci Sports Exerc* 1997; 29 **(Suppl)**: S1–S205.
6 Jacobs DR, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med Sci Sports Exerc* 1993; **25**: 81–91.
7 Ainsworth BE, Haskell WL, Leon AS, Jacobs DR, Montoye HJ, Sallis JF, Paffenbarger RS. Compendium of physical activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc* 1993; **25**: 71–80.
8 Williams E, Klesges RC, Hanson CL, Eck LH. A prospective study of the reliability and convergent validity of three physical activity measures in a field research trial. *J Clin Epidemiol* 1989; **42**: 1161–1170.
9 Chasan-Taber S, Rimm EB, Stampfer MJ, Spiegelman D, Colditz GA, Giovannucci E, Ascherio A, Willett WC. Reproducibility and validity of a self-administered physical activity questionnaire for male health professionals. *Epidemiology* 1996; **7**: 81–86.

10 Pols MA, Peeters PHM, Bueno-De-Mesquita BH, Ocke MC, Wentink CA, Kemper HC, Collette HJ. Validity and repeatability of a modified Baecke questionnaire on physical activity. *Int J Epidemiol* 1995; **24**: 381–388.
11 Drury TF (ed). *Assessing physical fitness and physical activity in population-based surveys*. National Center for Health Statistics: Hyattsville, 1989.
12 Taylor HL, Jacobs DR, Schucker B, et al. A questionnaire for the assessment of leisure time physical activities. *J Chronic Dis* 1978; **31**: 741–755.
13 Prentice AM (ed). *The doubly-labelled water method for measuring energy expenditure: Technical recommendations for use in humans. A consensus report by the International Dietary Energy Consultancy Working Group*. International Atomic Energy Agency: Vienna, 1990.
14 Black AE, Coward WA, Cole TJ, Prentice AM. Human energy expenditure in affluent societies: an analysis of 574 doubly-labelled water measurements. *Eur J Clin Nutr* 1996; **50**: 72–92.
15 Murgatroyd PR, Shetty PS, Prentice AM. Techniques for the measurement of human energy expenditure: a practical guide. *Int J Obes* 1993; **17**: 549–568.
16 Sequeira MM, Rickenbach M, Wietlisbach V, Tullen B, Schutz Y. Physical activity assessment using a pedometer and its comparison with a questionnaire in a large population survey. *Am J Epidemiol* 1995; **142**: 989–999.
17 Montoye HJ, Kemper HCG, Saris WHM, Washburn RA (eds). *Measuring Physical Activity and Energy Expenditure*. Human Kinetics: Champaign, 1996.
18 Suzuki I, Kawakami N, Shimizu H. Accuracy of Calorie Counter method to assess daily energy expenditure and physical activities in athletes and nonathletes. *J Sports Med Phys Fitness* 1997; **37**: 131–136.
19 Melanson EL, Freedson PS. Validity of the Computer Science Applications, Inc. (CSA) activity monitor. *Med Sci Sports Exerc* 1994; **27**: 934–940.
20 Miller DJ, Freedson PS, Kline GM. Comparison of activity levels using the Caltrac accelerometer and five questionnaires. *Med Sci Sports Exerc* 1994; **26**: 376–382.
21 Taylor CB, Kraemer HC, Bragg DA, Miles LE, Rule B, Savin WM, DeBusk RF. A new system for long-term recording and processing of heart rate and physical activity in outpatients. *Comp Biomed Res* 1982; **15**: 7–17.
22 Haskell WL, Yee MC, Evans A, Irby PJ. Simultaneous measurement of heart rate and body motion to quantitate physical activity. *Med Sci Sports Exerc* 1993; **25**: 109–115.
23 Matthews CE, Freedson PS. Field trial of a three-dimensional activity monitor: comparison with self-report. *Med Sci Sports Exerc* 1995; **27**: 1071–1100.
24 Welk GJ, Corbin CB. The validity of the Tritrac-R3D Activity Monitor for the Assessment of Physical Activity in Children. *Res Q Exerc and Sport* 1995; **66**: 202–209.
25 Payne PR, Wheeler EF, Salvosa CB. Prediction of daily energy expenditure from average pulse rate. *Am J Clin Nutr* 1971; **24**: 1164–1170.
26 Kalkwarf HJ, Haas JD, Belko AZ, Roach RC, Roe DA. Accuracy of heart-rate monitoring and activity diaries for estimating energy expenditure. *Am J Clin Nutr* 1989; **49**: 37–43.
27 Christensen CC, Frey HMM, Foenstelien E, Aadland E, Refsum HE. A critical evaluation of energy expenditure estimates based on individual O2 consumption/heart rate curves and average daily heart rate. *Am J Clin Nutr* 1983; **37**: 468–472.
28 Schulz S, Westerterp KR, Bruck K. Comparison of energy expenditure by the doubly labeled water technique with energy intake, heart rate and activity recording in man. *Am J Clin Nutr* 1989; **49**: 1146–1154.
29 Davey-Smith G, Morris JN. Assessment of physical activity, and physical fitness, in population surveys. *J Epidemiol Commun Health* 1992: 89–91.
30 Spurr GB, Prentice AM, Murgatroyd PR, Goldberg GR, Reina JC, Christman NT. Energy expenditure from minute-by-minute heart-rate recording: comparison with indirect calorimetry. *Am J Clin Nutr* 1988; **48**: 552–559.
31 Wareham NJ, Hennings SJ, Prentice AM, Day NE. Feasibility of heart-rate monitoring to estimate total level and pattern of energy expenditure in a population-based epidemiological study: the Ely young cohort feasibility study 1994–5. *Br J Nutr* 1997 (in press).
32 Livingstone MBE, Coward WA, Prentice AM, Davies PSW, Strain JJ, McKenna PG, Mahoney CA, White JA, Stewart CM, Kerr M-JJ. Daily energy expenditure in free-living children: comparison of heart-rate monitoring with the doubly labeled water method. *Am J Clin Nutr* 1992; **56**: 343–352.
33 Livingstone MBE, Prentice AM, Coward WA, Ceesay SM, Strain JJ, McKenna PG, Nevin GB, Barker ME, Hickey RJ. Simultaneous measurement of free-living energy expenditure by the doubly labeled water method and heart-rate monitoring. *Am J Clin Nutr* 1990; **52**: 59–65.
34 Ceesay SM, Prentice AM, Day KC, Murgatroyd PR, Goldberg GR, Scott W. The use of heart rate monitoring in the estimation of energy expenditure: a validation study using indirect whole-body calorimetry. *Br J Nutr* 1989; **61**: 175–186.
35 Manson JE, Rimm EB, Stampfer MJ, Colditz GA, Willett WC, Krolewski AS, Rosner B, Hennekens CH, Speizer FE. Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. *Lancet* 1991; **338**: 774–778.
36 Helmrich SP, Ragland DR, Leung RW, Paffenbarger RS. Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. *N Engl J Med* 1991; **325**: 147–152.
37 Manson JE, Nathan DM, Krolewski AS, Stampfer MJ, Willett WC, Hennekens CH. A prospective study of exercise and incidence of diabetes among US male physicians. *JAMA* 1992; **268**: 63–67.
38 Perry IJ, Wannamethee SG, Walker MK, Thomson AG, Whincup PH. Prospective study of risk factors for development of non-insulin dependent diabetes in middle aged British men. *BMJ* 1995; **310**: 560–564.
39 Burchfiel CM, Sharp DS, Curb JD, Rodriguez BL, Hwang L-J, Marcus EB, Yano K. Physical activity and Incidence of Diabetes: The Honolulu Heart Program. *Am J Epidemiol* 1995; **141**: 360–368.
40 Lynch J, Helmrich S, Lakka T, Kaplan GA, Cohen RD, Salonen R, Salonen JT. Moderately intense physical activities and high levels of cardiorespiratory fit-ness reduce the risk of non-insulin-dependent diabetes mellitus in middle-aged men. *Arch Intern Med* 1996; **156**: 1307–1314.
41 Haapanen N, Milunpalo S, Vuori I, Oja P, Pasanen M. Association of Leisure Time Physical Activity with the Risk of Coronary Heart Disease, Hypertension and Diabetes in Middle-Aged Men and Women. *Int J Epidemiol* 1997; **26**: 739–747.
42 Washburn RA, Goldfield SRW, Smith KW, McKinlay JB. The validity of self-reported exercise-induced sweating as a measure of physical activity. *Am J Epidemiol* 1990; **132**: 107–113.
43 Siconolfi SF, Lasater TM, Snow RCK, Carleton RA. Self-reported physical activity compared with maximal oxygen uptake. *Am J Epidemiol* 1985; **122**: 101–105.
44 Paffenbarger RS, Blair SN, Lee I-M, Hyde RT. Measurement of physical activity to assess health effects in free-living populations. *Med Sci Sports Exerc* 1993; **25**: 60–70.
45 King H, Rewers M. Global estimates for prevalence of diabetes mellitus and impaired glucose tolerance in adults. *Diabetes Care* 1993; **16**: 157–177.
46 Kriska AM, Bennett PH. An epidemiological perspective of the relationship between physical activity and NIDDM: from activity assessment to intervention. *Diabetes Metab Rev* 1992; **8**: 355–372.

47 Prentice AM, Jebb SA. Obesity in Britain: gluttony or sloth? *BMJ* 1995; **311**: 437–439.
48 Centers for Disease Control. *1992 BRFSS Summary Prevalence Report*. US Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Chronic Disease Prevention and Health Promotion: Atlanta, 1992.
49 Colhoun H, Prescott-Clarke P (eds). *Health Survey for England 1994, vol 1: Findings*. HMSO: London, 1996.
50 Eriksson KF, Lindgarde F. Prevention of type 2 (non-insulin-dependent) diabetes mellitus by diet and physical exercise. The 6-year Malmo feasibility study. *Diabetologia* 1991; **34**: 891–898.
51 Health Education Authority. *Physical activity Update*. Health Education Authority: London, 1995.
52 Killoran AJ, Fentem P, Caspersen C (eds). *Moving on: International Perspectives on promoting physical activity*. Health Education Authority: London, 1995.
53 Caspersen CJ. Physical activity epidemiology: concepts, methods, and applications to exercise science. *Exerc Sports Sci Rev* 1989; **17**: 423–473.