

REVIEW

Limits to the measurement of habitual physical activity by questionnaires

R J Shephard

Despite extensive use over 40 years, physical activity questionnaires still show limited reliability and validity. Measurements have value in indicating conditions where an increase in physical activity would be beneficial and in monitoring changes in population activity. However, attempts at detailed interpretation in terms of exercise dosage and the extent of resulting health benefits seem premature. Such usage may become possible through the development of standardised instruments that will record the low intensity activities typical of sedentary societies, and will ascribe consistent biological meaning to terms such as light, moderate, and heavy exercise.

Accurate measurement of habitual physical activity is fundamental to both the epidemiological study of relations between physical activity and health^{1,2} and the recommendation of an appropriate pattern of physical activity to maintain good health.³ If small numbers of subjects are to be studied, activity patterns can be determined in many different ways, including direct calorimetry, the ingestion of doubly labeled water, the use of motion sensors, accelerometers, heart rate recorders, or oxygen consumption meters, direct observation of movements by a trained observer, or assessments of food intake.⁴⁻¹⁰ However, epidemiological studies are often concerned with rare events, and the physical activity of large populations must then be categorised in order to draw significant conclusions. In the past, reported occupation has been successfully used to classify level of physical activity,^{11,12} but mechanisation, automation, and the skills of the ergonomist¹³ have together reduced the energy cost of most jobs to a point where an occupation based categorisation of activity is no longer of great value. A questionnaire has thus become the only feasible method of assessing habitual physical activity in large populations.^{5,7,14,15}

This review of recent literature looks critically at various problems with questionnaire assessment of the type, intensity, frequency, and duration of physical activity, and the environment in which it is performed. It explores how each of these variables may best be measured, summarising information on the reliability and validity of different types of instrument. It stresses that interrelations among the scores from rival questionnaires, and relations between individual scores and other measures of health status are at

best moderate.^{14,16} It also notes that attempts to interpret the data quantitatively can result in quite large absolute errors.¹⁷ Finally, some practical lessons are drawn for the epidemiologist and those prescribing physical activity.

TYPES OF PHYSICAL ACTIVITY

After many years of confusion, a consensus has now been reached on the definitions of physical activity, exercise, sport, recreation, occupational activity, and household chores.¹⁸⁻²⁰ Physical activity comprises all types of muscular activity that increase energy expenditure substantially. Exercise is a regular and structured subset of physical activity, performed deliberately and with a specific purpose such as preparation for athletic competition or the improvement of some aspect of health. Concepts of sport still differ between North America and some European countries. In North America, sport necessarily implies an activity that involves competition, whereas in Europe it may include recreational activities such as walking or hiking. Some forms of sport—for example, fishing and motor racing—do not involve a great deal of physical activity, and others, such as ice hockey and baseball, may become a job rather than a voluntary form of activity. Competition can be a source of motivation and self esteem for those who are successful, but it also increases the risk of both cardiovascular^{21,22} and musculoskeletal injury.²³ Recreational activity varies widely in its intensity, and the participant may attach value to the environment in which it is pursued. The workplace was once a major source of weekly energy expenditure, but in developed countries it has become a progressively less important component for most people.¹³ Household and other chores are a significant but sometimes largely overlooked component of the total weekly energy expenditure, particularly in full time caregivers.^{24,25}

For some purposes, it is useful to distinguish the predominant type of activity,²⁶ and indeed some surveys have distinguished occupational from leisure activity,²⁷ or occupational, sport, and leisure activity.²⁸ However, most aspects of health depend on the total amount of activity that is performed, and provided that the individual types of activity are all recognised and included in a global assessment, primary interest attaches to this overall score.

PATTERNS OF PHYSICAL ACTIVITY

Surveys have commonly focused on the intensity, frequency, duration, and total amount of physical activity performed. The relative proportions of

Correspondence to:
Professor Shephard, PO
Box 521, Brackendale, BC
V0N 1H0, Canada;
royjshep@shaw.ca

Accepted 9 August 2002

aerobic and resistance activity and the environmental context have attracted less attention.

Intensity

The intensity of physical activity may be expressed in absolute terms, as an absolute expenditure relative to body mass or resting metabolism, or as a value relative to peak performance.²⁹ From the viewpoint of physical conditioning, it has long been asserted that the last is the most important characteristic.³⁰ In the case of aerobic training, attempts have thus been made to express data as percentages of maximal oxygen intake, as fractions of the heart rate reserve, or most recently as fractions of the oxygen transport reserve.²⁰ Likewise, resistance activity has commonly been expressed as a fraction of the one repetition maximum contraction force for a given muscle group.²⁰

In support of the use of relative units, the physiological response to any given absolute intensity of effort appears to be greater in those who are unfit, or who have low initial levels of cardiorespiratory and muscular function because of aging.^{29, 30} The influence of relative intensity seems logical and is well supported by experimental data.^{30, 31} Nevertheless, one important remaining issue is the contribution to the observed relation of a reversion of data towards the true mean value as subjects who have been sorted in terms of their fitness are reassessed.³¹

Despite strong empirical arguments for expressing data relative to peak ability, many authors still focus on the absolute intensity of effort. For instance, subjects are asked to describe a typical speed of walking, jogging, or cycling. Reference tables are then used to convert such information into an approximate estimate of energy expenditure (kJ/min), oxygen consumption (litres/min or ml/min per kg), or metabolic activity relative to resting conditions (METs).^{4, 32, 33} Unfortunately, most standard compendia of metabolic costs are based on data for young adults, and they tend to overestimate the intensity of activity in middle aged and older people.^{34, 35} Translation of an absolute rate of energy expenditures to an estimate of relative intensity is possible only if some estimate of the person's maximal performance is available. Activities are often classed simply as light, moderate, hard, and very hard, although it is not always appreciated that the energy expenditures corresponding to such a perception depend on the duration of activity and the age and fitness of the person.¹⁸ For example, a young adult is likely to perceive a 20 minute bout of exercise that demands 50% of maximal oxygen intake as quite light activity, whereas if an older person is asked to maintain a 50% effort over an eight hour working day, the task is regarded as very hard.¹⁸

Some authors fail to distinguish clearly between gross and net energy expenditures. Both conditioning effects and the impact on metabolic problems such as obesity and diabetes mellitus depend on the net increase over resting energy expenditures. Further, as recently pointed out, the confounding of gross with net costs can create an apparent threshold energy expenditure of 2 MJ/week for health benefits.¹⁷

Frequency

Frequency is usually expressed as the number of times a given activity is performed a week. In countries with large seasonal extremes of temperature, both overall participation and the frequency of specific activities differ widely between summer and winter months.⁴ A second important aspect of frequency is whether a person takes all of a day's activity in a single session, or whether the activity is split into several smaller parts.³⁶ The latter approach is likely to encourage compliance in people who begin a prescribed exercise programme with a very low initial fitness. One report found a trend to a greater conditioning response with undivided sessions,³⁷ but several investigators found that, after adjustment to a common total

energy expenditure, gains of fitness were similar with single and divided sessions.^{36, 38} The influence of divided sessions on health outcomes remains to be determined.³⁶

Duration and amount

Information on the duration of individual exercise sessions may be combined with frequency data to indicate the total number of minutes of activity accumulated—for example, in a typical recent week. If the absolute intensity of effort has also been estimated, approximate figures can be cited for the corresponding gross or net weekly energy expenditure, expressed in kJ or MET.min.³⁹ Notice, however, that if health benefit is obtained from prolonged bouts of low intensity activity, much of the supposed “threshold” of gross energy expenditure reflects resting metabolism, and equal health benefit is likely if a lesser gross expenditure is developed at a higher average intensity of effort.¹⁷

Aerobic versus resistance activity

Current exercise recommendations call for an appropriate balance of aerobic and resistance activity,^{40–42} with sufficient weight bearing activity to enhance bone health. The impact of such activity on physical condition and health depends on the muscle groups that are exercised, the forces developed as fractions of maximal force for those muscle groups, the number of repetitions of contractions per set, the number of sets performed, and recovery intervals between sets.⁴³

Environmental context

Environment is a rarely noted aspect of physical activity. However, activity may be performed indoors or outdoors, in air or in water, under the relaxed conditions of a beautiful resort, or during a brief lunch break in a noisy gymnasium. The environmental conditions may be hot, with a high humidity and much radiant heating, or cold with a high wind chill factor. The ground surface may be smooth, rough, snow covered, or icy. Many of these factors alter the energy cost of a given activity,¹ and some (such as heat stress) interact with training responses⁴⁴ and changes in immune function.⁴⁵

Environment also influences the extent of ultraviolet exposure, the liability to heat stress or cold injury, the risk of musculoskeletal injury,^{22, 46} and the impact of a given bout of physical activity on psychological health (particularly the stimulation received by a person in boring work, or the relaxation experienced by a person who is overstressed).⁴⁷ Nevertheless, it remains arguable that, if the intensity, frequency, and duration of activity are established reliably, the environment has relatively little influence on many aspects of the aerobic response.

IMPLICATIONS FOR QUESTIONNAIRE DESIGN AND ANALYSIS

Basic issues

Questionnaires vary greatly in their detail, the period surveyed, and the extent of supervision of respondents.^{7, 48} Some investigators have made only a simple global classification of subjects—for example, active versus inactive—or have asked only a very few simple questions,^{49–51} presenting data as a three to five category scale, an arbitrary summary index (exercise units), or a simple continuous variable (for example, MET.min of activity per week). Care must be taken in interpreting ordinal scales, because intercategory increments of energy expenditure or total activity may not be uniform. Other studies have used lengthy forms that require up to an hour to complete, often with assistance from a trained observer.^{15, 52} Intercorrelations among the scores obtained from various complex questionnaires, and the correlations between such scores and assessments based on very few items, are often very low (0.14–0.41).^{53, 54} Further, perhaps because

subjects become bored and/or confused by lengthy instruments, some of the highest coefficients of reliability and validity are seen for simple questionnaires.^{14–55} A comparison of the Baecke and Tecumseh questionnaires concluded that the former yielded superior results because it was simpler.⁵⁶ Likewise, test/retest correlations were 0.81 for the very simple Godin questionnaire and 0.93 for a simple activity rating,⁵⁵ and the Godin questionnaire also fared better than more complex instruments relative to a Caltrac motion assessment of concurrent validity.

Questionnaires may examine activities during the past one to seven days,^{49–57–58} through the last month^{32–59} to (in some instances) an entire lifetime.^{60–63} If the recording period is less than one week, care must be taken to include both weekday and weekend activities.^{64–65} Population sampling should also be dispersed to take account of seasonal variations in activity patterns.⁶⁶ Questionnaire responses depend on the perception, encoding, storage, and retrieval of information about previous physical activity; answers depend on the subject's age and the context of questioning.^{67–68} Because of limitations in human memory, the reliability of information generally decreases with the length of the period surveyed, and it is best to keep the reporting interval relatively short (no longer than three months⁴); however, in advanced age, long term memory may be better preserved than recent recollection of activity patterns. The accuracy of responses may be helped by asking questions about a specific time of day (for example, "what do you do after supper?"⁶⁹). Interval response options—for example, less than twice a week, two to three times a week, four to five times a week—may elicit a higher apparent frequency of physical activity than open ended questions.⁷⁰

Montoye *et al*⁷¹ insisted that subjective responses to the detailed Tecumseh questionnaire required considerable interpretation. There was good correspondence between the activity ratings of three judges, but self reports were unsatisfactory unless combined with an interview.⁷² Another study found underestimation of activity by self report in military officers, but a better agreement between self reports and interviewers in ordinary working men.⁵¹ However, the mediation of an interviewer has sometimes had surprisingly little impact on the total amount of activity reported.⁷³ In one study, self reports showed somewhat less leisure activity than an interviewer assessment, but the one year activity scores reported by the subject and an interviewer correlated closely ($r = 0.83$).⁷⁴ Interobserver ($r = 0.42–0.99$) and intraobserver ($r = 0.56–0.96$) inconsistencies were also noted by these investigators.⁶³

Questionnaire responses can be influenced considerably by cultural factors, in part because the content of reported activities differs from one country to another, and in part because the manner of answering questions shows a cultural bias.⁴ Difficulties are particularly likely if a questionnaire has been translated into another language.⁷⁵ Respondents may also be influenced by the social desirability of reporting particular behaviours. In general, people tend to overreport physical activity and underestimate sedentary pursuits such as watching television.^{76–77} Sims and associates noted specifically that people who had been encouraged to exercise reported a greater volume of physical activity than could be confirmed by heart rate data.⁷⁸

In young children and those with mental impairments, attempts must be made to deduce physical activity patterns from the questioning of guardians, but such estimates have poor reliability and validity.⁷⁹ In the elderly, other problems may arise from impairments of vision and hearing, and disturbances of cognition.⁸⁰

Types of activity

The number and type of activities reported can be augmented substantially by the use of either cue cards⁴ or leading

questions on the part of an interviewer. A Swiss study suggested the value of listing 70 activities that together accounted for 95% of the weekly energy expenditure; if subjects indicated the number of days each activity was performed, and a typical duration for the activity, the resulting score was held to correlate well with estimates based on a heart rate monitor ($r = 0.76$).⁸¹ However, it remains uncertain whether such prompting recalls activities that have been performed very rarely, or whether they are important components of the total picture that the subject has inadvertently overlooked.

If interest is focused on specific activities, difficulty may arise because subjects obtain most of their weekly physical activity from items that are not listed on the cue cards. There have been suggestions that this approach may underestimate the volume of activity performed by full time caregivers^{24–25} and the elderly.⁸² However, a recent study based on doubly labelled water found that elderly women tended to overestimate their involvement in high scoring components of housework.⁸³

Intensity

Questionnaires often express the intensity of physical activity semantically, using a Likert-type scale. Unfortunately, perceptions of the intensity of any stimulus depend on the experience and the stoicism of the person concerned.⁸⁴ Some people are particularly prone to report symptoms.⁸⁵ Reporting may also be influenced by the perceived desirability of a given response.⁷⁶ In the case of sport, additional information on the intensity of activity may be derived from the level of competition, the number of training sessions a week, and the time required to perform a standard task such as swimming four lengths of a 25 m pool. At work, where all intensities of effort are relatively low, distinction may be drawn between portions of the day spent sitting, standing, walking, and lifting or carrying.⁸⁶

There have been attempts to anchor semantic descriptions of exercise intensity in physiological terms, as with the original Borg scale, where each unit of perceived intensity was intended to correspond to a 10 beats/min increase in heart rate.⁸⁷ Over a typical 15–30 minute bout of endurance exercise, the average person will perceive a task in the aerobic training zone as moderately hard (a Borg rating of 12–14 units,⁸⁷ corresponding in a middle aged adult to a heart rate of 120–140 beats/min). Other potential anchors of intensity include "exercise sufficient to induce moderate sweating"⁴⁹ or "causing sufficient breathlessness to limit conversation".⁸⁸ However, such descriptions at best distinguish light from vigorous effort. They are again somewhat vulnerable to differences in symptom reporting,⁸⁵ and in the case of sweating are affected by environmental temperatures.

A further issue is the probable need to measure very low levels of physical activity. Unfortunately, many questionnaires suffer from floor effects.⁸⁹ For example, one widely used seven day recall instrument does not take account of activities that are less intense than brisk walking, or that have a duration of less than ten minutes.⁹⁰ The shape of the dose/response curve remains unclear,⁹¹ but some recent research suggests that, particularly in the frail elderly and those who are extremely sedentary, health advantages may accrue from very low levels of physical activity that are unlikely to induce breathlessness, sweating, or an increase in aerobic fitness.^{92–94}

Many questionnaires focus on the absolute rather than the relative intensity of individual physical activities. For instance, subjects are asked to specify a typical speed of walking, jogging, or cycling.⁴ Using a table of energy costs,³² many (but not all) reported activities can be converted into an approximate estimate of the rate of energy expenditure (kJ/min), an intensity of metabolic activity relative to resting conditions (METs), or an oxygen consumption (ml/min/per kg).

Nevertheless, there are substantial interindividual and intra-individual variations in the energy cost of various activities, depending on the subject's age, sex, body mass, skill, and level of fatigue.^{49–96} For example, the pace of walking differs considerably between those who are undertaking the activities of daily living and those who are performing deliberate exercise.² Moreover, the costs of some activities are either unknown, or have changed since the data were first collected. Finally, some authors have translated data to MET values by assuming a standard value of 4.19 kJ/min (1 kcal/min) for basal metabolism; in fact, values vary with age, sex, and body surface from about 3.47 to 2.55 kJ/min per m².

Measures of absolute and relative intensity of effort may give widely differing estimates of the prevalence of adequate physical activity.⁷⁰ Translation of absolute data to a relative intensity of activity is possible only if the subject's maximal oxygen intake is known.⁹⁶ Questionnaires do not normally provide such information. There have been suggestions that subjects have a fairly clear perception of their physical fitness^{97–98} and that moderately accurate predictions of fitness can be made from age, body mass, skinfold thicknesses, and global assessments of habitual physical activity without engaging subjects in an exercise test.^{97–102} Nevertheless, the generality of such prediction equations is questionable,¹⁰³ and the confidence limits are so broad that it is difficult even to categorise a person's fitness status.¹⁰⁴

Some analyses have apparently confounded the intensity of energy expenditure with the total quantity of energy expended a week. For example, one recent review described those with an energy expenditure of >20 MET.h/week as "highly active".³¹ However, this expenditure could have been reached through 13 h/week of occupational activity at an intensity of only 1.5 METs. Increases in the total weekly energy expenditure that are achieved by moderate or low intensities of activity can be important for some aspects of metabolic health, although data summarised as MET.h or kJ/week cannot answer questions about the importance of a given absolute or relative intensity of physical activity to the prevention of ischaemic heart disease.

Frequency

The frequency of activity is usually reported as times per week or times per month. This may be a reasonable approach when making an overall assessment of habitual activity—for example, the number of sessions of sweat inducing activity of 20 minutes duration or longer.⁴⁹ However, if such an assessment is applied repeatedly to a wide range of individual activities such as walking, running, cycling, and swimming, subjects are liable to overestimate the total hours of activity that they perform in a week.

The frequency of many activities varies substantially, even over an interval of a few months, and unfortunately respondents are liable to indicate their highest recent or their desired rather than their true average frequency of participation. Such problems are compounded by seasonal changes in activity patterns.⁴ However, if a large population is to be examined, true population averages can be approximated by dispersing questionnaire assessments over an entire calendar year.

In view of recent research on the value of split exercise sessions for frail patients,³⁶ it may be useful to keep track of divided sessions during a given day, particularly in those with low levels of fitness.

Duration

The duration of some types of activity tends to be over-reported, making it necessary to adjust reported data substantially in order to limit the total length of a subject's day to 24 hours.⁴ In addition to problems of exaggeration, the indicated minutes of attendance at a sports club may include time devoted to changing, refreshment, and socialising.¹⁰⁵

Some of the largest overestimates of exercise duration come from the school gymnasium, where the major fraction of a 30 or 40 minute physical education class may be spent in listening to instructions and awaiting a turn to use a particular item of equipment.¹⁰⁶ Problems of underreporting can also arise through failure to take account of brief periods of activity encountered during some forms of everyday activity.⁸⁹

In young children, questionnaire assessments of physical activity are greatly complicated by their propensity for repeated brief bouts of vigorous physical activity.¹⁰⁷

Timing and overall duration of activity

Given that people do not maintain a consistent exercise behaviour throughout their lifetime, it may be important to ascertain when activity has been performed. In terms of the heart, current physical activity seems the most important determinant of both fitness and health. Little or no benefit is found from former athletic¹⁰⁸ or leisure activity.¹⁰⁹ But in terms of the prevention of osteoporosis, the critical factor may be the maximisation of bone mass during early adult life,^{110–111} and the key to prevention of some neoplasms may lie in adequate activity during adolescence, when cell division is at a maximum.¹¹² Unfortunately, attempts to determine either the amount or the intensity of physical activity performed many years previously have only limited reliability and validity (see below).

Amount of physical activity

The impact of physical activity on certain metabolic variables such as obesity, the risk of diabetes mellitus, and hypercholesterolaemia seems to depend mainly on the total amount of energy expended, and many reports have summarised activity levels in such terms.¹² However, increases in energy expenditure are necessarily the product of the net intensity and the duration of activity, and commonly the two variables are confounded. A large energy expenditure is usually accumulated because a person chooses to exercise at a relatively high intensity of effort. At least one analysis has suggested that energy expenditures accumulated in non-vigorous physical activity do not influence longevity.¹¹³

Aerobic versus resistance activity

Few questionnaires have addressed the issue of the relative proportions of aerobic and resistance activity.³ Information on resistance exercise is likely to be available if the subject has undertaken some type of circuit training, but the extent of such activity is very difficult to determine if the main type of activity is the performance of normal daily activities. Investigators are currently exploring potential questionnaires that can assess the extent of resistance activity.¹¹⁴

Environmental issues

To my knowledge, none of the existing questionnaires of personal physical activity habits explore the type of environment in which an individual normally undertakes physical activity. The type of environment has particular importance in the contexts of motivation and the psychological benefits of physical activity.

RELIABILITY, VALIDITY, AND SENSITIVITY OF MEASURING INSTRUMENTS

A number of questionnaires were used quite widely before issues of their reliability and validity had been addressed. The current number of questionnaires suggests that many do not yield either reliable or valid information. Estimates of the prevalence of limited physical activity among women of child bearing age in the United States have ranged from 3.9% to 39.0%, using questions from three surveys conducted by the National Center for Health Statistics.¹¹⁵ Likewise, in the behavioural risk factor surveillance system, the prevalence of

moderate activity as assessed by differing algorithms ranged from 20% to 38%.¹¹⁶ The proportion of the population of the United States who appear to meet current fitness guidelines varies from 32% to 59%, depending on the test instrument and the scoring protocol used; the proportion meeting health related guidelines varies even more widely, from 4% to 70%.⁷⁰ Factors contributing to this wide variability include not only personal characteristics (age, sex, and socioeconomic status) but also the use of prompting cards and/or questions, and the number of items included in the estimate—for example, leisure activity versus leisure + occupation + household + transportation.

Reliability

The reliability of a questionnaire reflects its ability to yield the same result if it is applied on a second occasion. If the test is administered by an observer, variance due to errors of subject reporting and true changes in activity patterns between the two assessments are compounded by interobserver and/or intraobserver errors.⁶⁵ Looking retrospectively at the total activity accumulated over 20 years, one report noted an inter-observer reliability coefficient of about 0.90, and an intra-observer coefficient (on an older group of subjects than the interobserver study) of 0.70.⁶³

Appropriate statistical methods must be used in assessing the reliability and validity of questionnaire responses.¹¹⁷ Several authors have used χ^2 statistics.¹¹⁸ Booth and associates¹¹⁹ evaluated the World Health Organisation Health Behaviour in Schoolchildren (WHO HBSC) questionnaire, showing a 70% agreement between two way classifications of 13 and 15 year old children (active/insufficiently active) after an interval of two weeks. Other statistically acceptable alternatives include calculations of the coefficient of variation in response (SD/mean) or the coefficient of repeatability (= 2SD). A small scale analysis of minutes per week of moderate or greater leisure and occupational activity found respective coefficients of repeatability of 29.3 and 54.6 minutes on totals of about 150 and 450 minutes.¹²⁰

More commonly, reliability has been evaluated in terms of test/retest intraclass correlation coefficients. This is a less desirable approach,¹²¹ in part because the magnitude of correlations is influenced by the extent of interindividual variance within the data set. If the samples evaluated are uniformly sedentary, low coefficients of reliability and validity are to be expected. Of potentially recorded characteristics (intensity, frequency, and duration), intensity seems to be the least reliably reported.¹²² Problems arise from differing individual perceptions of a given absolute or relative intensity of effort, and a lack of agreement on MET values corresponding to vigorous, moderate, and light activity. The reliability of responses also varies with the interval between tests. Two week test/retest observations on the simple questionnaire of Godin and Shephard⁴⁹ found an intraclass correlation coefficient of 0.94 for reports of strenuous activity (estimated intensity 9 METs), falling to only 0.46 for moderate activity (5 METs) and 0.48 for light activity (3 METs).

Most authors have looked at indices of total activity. Reliability diminishes with the length of the recall period. Again, this has been assessed by test/retest correlations. Lamb and Brodie¹⁰² found a two week coefficient of 0.86, and the five week coefficient for the Minnesota leisure time physical activity questionnaire was 0.88.¹²³ Studies on the college alumnus questionnaire found *r* values of 0.72 at one month, falling to 0.3–0.4 over 8–12 months.^{14 32 124 125} Other authors have reported coefficients of a similar order: 0.58–0.67 for the community health activities model program for seniors (CHAMPS) physical activity questionnaire over a six month interval,¹²⁶ 0.55 for adolescents over one year,¹²⁷ and 0.59 for a two to three year recall in the coronary artery risk development in young adults (CARDIA) study.¹²⁸

Lack of reliability is due in part to seasonal and/or temporal variations in physical activity patterns, but shortcomings of human memory are also an important problem. Thus, questionnaire responses show a variation of 50% or more, even if one year activity patterns are reassessed after an interval of a few days. The problem is particularly acute if intensities of effort are low.¹²⁹

The reliability of absolute scores has received little attention, although there have been suggestions that, if a questionnaire is completed on several occasions, subjects become less precise in their responses, and intraindividual variations in reported activity diminish.¹⁰⁵

Validity

Quite a number of investigators have limited their questionnaire evaluations to an examination of reliability, neglecting the more important issues of validity and sensitivity of response.¹⁵ Validity has a number of components (content, predictive, concurrent, and construct).¹³⁰ Physical activity questionnaires should ideally be validated in terms of their criterion validity (a combination of predictive and concurrent validity, indicating the correspondence of scores to a more precise assessment of the characteristic of interest—for example, the total volume of physical activity performed). However, given the absence of any widely accepted criterion of physical activity,^{6 28} reliance has usually been placed on construct validation against other observations that are linked with physical activity. One analysis also looked at an expression of concurrent validity, comparing reported physical activity with the stage of change in exercise behaviour.¹²⁰

Measurements of energy expenditure using doubly labelled water are commonly accepted as the optimum in construct validation.¹³¹ The within subject variation for this technique (analytical plus biological variation) is about 8%.¹³² However, the necessary analyses are costly (as much as US\$600 for a single measurement), and the data at best provide a two week average of energy expenditure. Further, the energy cost of many activities varies substantially between people, and in some disease states data interpretation may be complicated by an increase in basal metabolism.¹³³ Tests against doubly labelled water have yielded correlation coefficients of 0.68 for the Baecke total activity index, 0.57 for the sweat index from the Five City Project questionnaire, 0.64 for the Tecumseh estimate of total energy expenditure,¹³⁴ and 0.79 in men and 0.68 in women for the physical activity scale for the elderly.⁸³

Some form of motion sensor such as a pedometer or accelerometer¹³⁵ has provided a second, cheaper construct. Such devices tend to underestimate walking and overestimate jogging activity, also failing to detect arm movements, resistance exercise, and the performance of external work.^{136 137} Motion sensor scores may also show only weak relations to maximal oxygen intake.^{138 139} A Dutch questionnaire showed correlations of 0.78 and 0.73 with 24 hour activity recall and pedometer measurements respectively. Two thirds of subjects were assigned to the corresponding activity tertile by each of these methods.¹⁴⁰ Other authors have found coefficients of around 0.70. Seven day scores for a Japanese pedometer showed correlations of 0.68–0.69 with questionnaire scores, the latter overestimating total energy intake by 4.5% in men, but not in women.¹⁴¹ Likewise, the correlation with the Minnesota leisure time physical activity questionnaire was 0.69 when a motion sensor was waist mounted, although it fell to 0.43 when the device was attached to the leg.¹⁴² Other investigators have found weaker correlations, possibly because most of their subjects were sedentary: one week activity scores had a correlation of 0.56 with Caltrac motion sensor values,¹²⁰ the coefficient relating three day portable accelerometer scores to the physical activity scale for the elderly was 0.49,¹⁴³ the Baecke and pre-EPIC questionnaires showed correlations of only 0.22 with Caltrac motion sensor

scores in elderly women,¹⁴⁴ and correlations between the Tecumseh community questionnaire and triaxial accelerometer scores were in the range 0.26–0.47.⁵⁶ In patients in which the intensity of activity is limited by chronic disease, the correlation with accelerometer scores or doubly labelled water measurements may become non-significant—for example, $r = 0.14$ in chronic obstructive pulmonary disease¹⁴⁵ and $r = 0.057$ in peripheral vascular disease.¹⁴⁶

Other approaches to validation have included comparisons with information obtained from heart rate records, exercise logs or diaries, 24 hour activity recalls, fitness scores, food consumption, and health outcomes.¹⁵ Significant correlations with aerobic fitness should be observed only for vigorous, sweat producing activity.¹⁴⁷ In keeping with this expectation, correlation coefficients have been largest for the hardest forms of activity.^{14 124 129 148} Testing correlations with physical work capacity, one study found a correlation of 0.55 for “very hard” leisure time activity and 0.48 for “hard” activity, but much weaker correlations for less intense activity.¹⁰² A second report examined correlations between exercise test scores and responses to the Minnesota leisure time physical activity questionnaire in Spanish women; correlations fell from 0.51 for heavy activity to 0.13 for moderate and 0.02 for light activity.¹⁴⁹

Harada and associates³⁴ found correlations of 0.44–0.68 between scores on three types of activity questionnaire (CHAMPS questionnaire, physical activity survey for the elderly, and the Yale physical activity survey) and performance based tests of lower body functioning and endurance. In a relatively active population, an overall index derived from the Harvard alumni questionnaire showed a correlation of 0.52 with aerobic power, although two of the three elements in this index (stepping and walking scores) showed almost zero correlation with aerobic performance (respective r values of 0.02 and 0.01).¹⁵⁰ The correlation between scores on the physical activity scale for the elderly and peak oxygen intake was only 0.20.¹⁴³ Likewise, an interviewer administered seven day activity recall showed very low correlations with estimated maximal oxygen intake (0.34), resting heart rate (−0.09), and body mass index (−0.23).¹⁵¹ Correlations with treadmill measurements of maximal oxygen intake were 0.31 for an index based on running, walking, and jogging, 0.35 for a question based on frequency of sweating,⁵⁸ and 0.29 for Paffenbarger's leisure time activity index.¹⁵² Other correlations were for treadmill run times (0.41),¹⁵³ for physical work capacity at a heart rate of 150 beats/min (0.08¹⁵⁴ and 0.10¹⁵⁵), for 1.6 km run times (−0.37),¹²⁷ for submaximal treadmill scores (0.13),⁷² and for total energy intake (−0.10,⁵³ 0.27,¹⁵⁶ or 0.02⁵⁷). Nevertheless, it was possible to account for some 75% of the variance in accelerometer scores of patients with end stage renal disease using a combination of the physical activity scale for the elderly and the human activity profile.¹⁵⁸

Godin and Shephard⁴⁹ developed discriminant functions to predict two way classifications of aerobic fitness and body fat content from responses to a simple activity questionnaire. Respective κ values were 0.30 and 0.17, with 69% and 66% of subjects being classified correctly.

Often, investigators have accepted validity coefficients of 0.3–0.5 relative to other direct or indirect measures of physical activity and energy expenditure. Thus Bairey-Merz and associates¹⁵⁹ concluded that the Duke activity status index was “a reasonable correlate of functional capacity,” given a coefficient of 0.31—that is, it described 9.6% of the total variance in functional capacity. A two year trial on nurses found “reasonably valid” measures: a test/retest correlation of 0.59, and a correlation of 0.60 between diary and questionnaire data.¹⁶⁰ A third report suggested that a questionnaire gave a “reasonable” estimate of physical activity over the past year, even though scores showed no significant relations to physical fitness or body mass index.¹⁶¹

Even in studies in which correlation coefficients have been relatively high, absolute estimates of physical activity have

shown large errors. Thus, one comparison between the college alumnus questionnaire and pedometer scores gave respective estimates of daily walking distance as 2.3 (1.6) and 6.7 (2.6) km.¹⁶² The gross energy cost of stair stepping was also only 50% of values assumed in the questionnaire analyses.¹⁶³ Nevertheless, the total physical activity (MET.min/week) indicated by the questionnaire was only a third of values found in a 48 hour physical activity log.¹²⁴ Attempts to recall activity that had previously been assessed 11 years ago also led to a 41% increase in the estimated total weekly energy expenditures.¹⁶⁴

Sensitivity

An effective questionnaire must be not only reliable and valid, but also sufficiently sensitive to detect relevant activity related differences in health status and programme related changes in patterns of habitual physical activity.¹⁶⁵ In some studies, in which little effect of a programme has been seen, it is difficult to be certain whether the problem lies with the programme or the measure of habitual physical activity.¹⁶⁶

In general, questionnaires seem to be less sensitive than more objective instruments such as accelerometers. Thus, accelerometer scores suggested that patients with peripheral arterial disease had only 46% of the energy expenditure of controls ($p < 0.001$), whereas applications of the health interview survey and the Stanford seven day activity recall to the same subjects suggested values that were 73% ($p = 0.128$) and 98% ($p = 0.454$) of controls respectively.¹⁶⁷ In Britain, civil servants classed as active had a significantly greater daily food intake than those classed as inactive.¹⁵⁶ Likewise, there were 12–23% differences of shuttle run score between “active” and “inadequately active” adolescents, using the WHO HBSC questionnaire.¹¹⁹

Blair and associates⁹⁰ tested the ability of a seven day activity recall to detect associations between changes in energy expenditure and gains of fitness over a 12 month trial. They found correlations of 0.33 for maximal oxygen intake, −0.50 for body fatness, and 0.32 for high density lipoprotein cholesterol.⁹⁰

Small to moderate size effects (0.38–0.64) were noted when one questionnaire was used to evaluate the impact of a six month programme promoting physical activity.¹²⁶ A much larger effect size of 1.68 was shown when pedometers were used to evaluate a four week programme for patients with type 2 diabetes mellitus.¹⁶⁸

CONCLUSIONS AND IMPLICATIONS FOR EPIDEMIOLOGY AND EXERCISE PRESCRIPTION

What practical conclusions can the epidemiologist and those formulating exercise prescriptions draw from this review? Irrespective of the questionnaire chosen, the data probably have limited reliability and validity relative to a laboratory measure of physical activity. If, as is commonly the case,¹² the need is to calculate risk ratios for two or three different volumes of habitual physical activity—for example, high, moderate, and low—the use of a large number of subjects commonly reduces problems resulting from imprecise classification and allows demonstration of activity related benefits of reduced morbidity and mortality and enhanced health. Nevertheless, misclassification reduces the apparent magnitude of any benefits from physical activity. This is an important reason why a three level classification of aerobic fitness (assessed accurately in the laboratory) apparently has a larger influence on all cause mortality than a three level classification of habitual physical activity.⁹¹

If large populations are examined, categorisation may also show significant dose/response relations.^{26 29} However, the absolute energy expenditures corresponding to light, moderate, and vigorous effort remain unclear, and attempts to interpret questionnaire data in a quantitative sense are generally

Take home message

Despite extensive use over 40 years, physical activity questionnaires still show limited reliability and validity. Measurements have value in indicating conditions where an increase in physical activity would be beneficial, and in monitoring changes in population activity. However, attempts at detailed interpretation in terms of exercise dosage and the extent of resulting health benefits seem premature. Such usage may become possible through the development of standardised instruments that will record the low intensity activities typical of sedentary societies, and will ascribe consistent biological meaning to terms such as light, moderate and heavy exercise.

unwarranted. Plainly, there is need for international consensus on the wording of questionnaires and the methods of analysis and interpretation.^{169 170} Both data interpretation and comparisons between studies would be greatly facilitated if individual observers used a reference standard such as doubly labelled water or even heart rate recordings on a small sample of their subjects, to clarify the average energy expenditures equivalent to each of their activity categories.

The choice of questionnaire depends ultimately on the purpose of the investigator, and the available resources of time, funding, and skilled personnel. Nevertheless, for many purposes an accurate but simple classification of activity levels may be more appropriate than an attempt at estimating overall energy expenditures. Care must be taken to avoid bias when distinguishing between categories of activity intensity or volume. Cut point bias may be introduced because categorisation is adjusted to fit sample distribution—for example, the use of tertiles—or to maximise statistical significance.¹⁷¹ Particular difficulty is experienced in detecting and assessing low levels of physical activity, and given that this is the most prevalent form of activity in the general North American population, attention must be focused on developing better methods to assess low intensity effort.¹¹⁴

Measurement errors assume particular significance in discussions of dose/response relations²⁶ and recommendations of a minimum dose of physical activity to optimise population health. Many authors claim gains from quite moderate intensities of effort (as little as 40–50% of oxygen consumption or heart rate reserve).^{94 172} Others have argued that, after adjustment for total energy expenditure and other confounding variables, no significant benefit is obtained unless the intensity of effort exceeds an absolute level of 6 METs and the total energy expenditure exceeds 2 MJ/week.^{113 173 174} Nevertheless, proponents of high intensity exercise found a trend ($p < 0.07$) to benefit from moderate activity (intensity > 4 METs), and it seems likely that if their questionnaire classifications of exercise intensity had been more precise, this trend would have been significant.

Likewise, much, if not all, of the apparent energy expenditure threshold of 2.1–4.2 kJ/week¹⁷⁵ would disappear if account were taken of systematic errors in the estimation of energy costs and data were expressed as net rather than gross energy expenditures.¹⁷

Despite the problems outlined in this review, physical activity questionnaires have practical value in indicating conditions where an increase in physical activity would be beneficial and in monitoring changes in population activity. Attempts at more detailed interpretation in terms of exercise dosage and the extent of resulting health benefits seem premature. However, such usage may become possible through the development of standardised instruments that will record the low intensity activities typical of sedentary societies, and will ascribe consistent biological meaning to terms such as light, moderate, and heavy exercise.

REFERENCES

- 1 Bouchard C, Shephard RJ, Stephens T. *Physical activity, fitness and health*. Champaign, IL: Human Kinetics, 1994.
- 2 Macera CA, Powell KE. Population attributable risk: implications of physical activity dose. *Med Sci Sports Exerc* 2001;**33**(suppl 6):S635–9.
- 3 Bouchard C. Physical activity and health: introduction to the dose-response symposium. *Med Sci Sports Exerc* 2001;**33**(suppl 6):S347–50.
- 4 Shephard RJ. *Fitness of a nation: lessons from the Canada Fitness Survey*. Basel: S Karger, 1986.
- 5 Lamonte MJ, Ainsworth BE. Quantifying energy expenditure and physical activity in the context of dose response. *Med Sci Sports Exerc* 2001;**33**(suppl 6):S370–8.
- 6 LaPorte RE, Montoye HL, Caspersen CJ. Assessment of physical activity in epidemiological research: problems and prospects. *United States Public Health Report* 1985;**100**:131–46.
- 7 Montoye HJ, Kemper HCG, Saris WHM, et al. *Measuring physical activity and energy expenditure*. Champaign, IL: Human Kinetics, 1996.
- 8 Sirard JR, Pate RR. Physical activity assessment in children and adolescents. *Sports Med* 2001;**31**:439–54.
- 9 Trost SG. Objective measurements of physical activity in youth: current issues, future directions. *Exerc Sport Sci Rev* 2001;**29**:32–6.
- 10 Wilson PWF, Paffenbarger RS, Morris JN, et al. Assessment methods for physical activity and physical fitness in population studies: report of a NHLBI workshop. *Am Heart J* 1986;**111**:1177–92.
- 11 Powell KE, Thompson PD, Caspersen CJ, et al. Physical activity and the incidence of coronary heart disease. *Annu Rev Public Health* 1987;**8**:253–87.
- 12 Shephard RJ, Fitcher R. Physical activity and cancer: How may protection be maximized? *Critical Reviews of Oncogenesis* 1997;**8**:219–72.
- 13 Shephard RJ. Assessment of occupational fitness in the context of human rights. *Can J Sport Sci* 1990;**15**:89–95.
- 14 Jacobs DR, Ainsworth BE, Hartman TJ, et al. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med Sci Sports Exerc* 1993;**25**:81–91.
- 15 Vuillemin A. Revue des questionnaires d'évaluation de l'activité physique [A review of evaluation questionnaires for physical activity]. *Revue Epidemiol Santé Publique* 1998;**46**:49–55.
- 16 Kriska AM, Caspersen CJ. A collection of physical activity questionnaires for health-related research. *Med Sci Sports Exerc* 1997;**29**(suppl 6):S1–205.
- 17 Shephard RJ. How much physical activity is needed for good health. *Int J Sports Med* 1999;**20**:23–7.
- 18 Bouchard C, Shephard RJ. Physical activity, fitness, and health: the model and the key concepts. In: Bouchard C, Shephard RJ, Stephens T, eds. *Physical activity, fitness and health*. Champaign, IL: Human Kinetics, 1994:77–88.
- 19 Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise and physical fitness: definitions and distinctions for health-related research. *United States Public Health Reports* 1985;**100**:126–31.
- 20 Howley ET. Type of activity: resistance, aerobic and leisure versus occupational physical activity. *Med Sci Sports Exerc* 2001;**33**(suppl 6):S364–9.
- 21 Vuori I. Exercise and sudden cardiac death: effects of age and type of activity. *Sports Sci Rev* 1995;**4**:46–84.
- 22 Thompson PD, Fahrenbach MC. Risks of exercising: cardiovascular, including sudden death. In: Bouchard C, Shephard RJ, Stephens T, eds. *Physical activity, fitness and health*. Champaign, IL: Human Kinetics, 1994:1019–28.
- 23 Pate RR, Macera CA. Risks of exercising: musculoskeletal injuries. In: Bouchard C, Shephard RJ, Stephens T, eds. *Physical activity, fitness and health*. Champaign, IL: Human Kinetics, 1994:1008–18.
- 24 Blair SN, Kohl HW, Barlow CE. Physical activity, physical fitness, and all-cause mortality in women: do women need to be active? *J Am Coll Nutr* 1993;**12**:368–71.
- 25 Masse LC, Ainsworth BE, Tortolero S, et al. Measuring physical activity in midlife, older and minority women: issues from an expert panel. *J Women's Health* 1998;**7**:57–67.
- 26 Wareham NJ, Rennie KL. The assessment of physical activity in individuals and populations: why try to be more precise about how physical activity is assessed? *Int J Obes Metab Disord* 1998;**22**(suppl 2):S30–8.
- 27 Shapiro S, Weinblatt E, Frank CW, et al. The HIP study of incidence and prognosis of coronary heart disease: preliminary findings on incidence of myocardial infarction and angina. *Journal of Chronic Diseases* 1965;**18**:527–58.
- 28 Baecke JAH, Burema J, Fritjers JER. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutr* 1982;**36**:932–42.
- 29 Shephard RJ. Absolute versus relative intensity of physical activity in a dose-response context. *Med Sci Sports Exerc* 2001;**33**(suppl):S400–18.
- 30 Shephard RJ. Intensity, frequency and duration of exercise as determinants of the response to a training regimen. *Int Z Angew Physiol* 1968;**26**:272–8.
- 31 Shephard RJ, Hamm L, Mertens DJ, et al. Influence of initial fitness on the response to cardiac rehabilitation [abstract]. *Med Sci Sports Exerc* 2002;**34**:S181.
- 32 Ainsworth BE, Haskell WL, Leon AS, et al. Compendium of physical activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc* 1993;**25**:71–80.

- 33 **Passmore R**, Durnin JVGA. Human energy expenditure. *Physiol Rev* 1955;**35**:801-40.
- 34 **Harada ND**, Chiu V, King AC, *et al*. An evaluation of three self-report physical activity instruments for older adults. *Med Sci Sports Exerc* 2001;**33**:962-70.
- 35 **Rikli RE**. Reliability, validity and methodological issues in assessing physical activity in older adults. *Res Q* 2000;**71**:89-96.
- 36 **Hardman AE**. Issues of fractionalization of exercise (short vs long bouts). *Med Sci Sports Exerc* 2001;**33**(suppl 6):S421-8.
- 37 **DeBusk R**, Stenestrand U, Sheehan M, *et al*. Training effects of long versus short bouts of exercise in healthy subjects. *Am J Cardiol* 1990;**65**:1010-13.
- 38 **Lee I-M**, Sesso HD, Paffenbarger RS. Physical activity and coronary artery disease risk in men: does the duration of exercise episodes predict risk? *Circulation* 2000;**102**:981-6.
- 39 **Paffenbarger RS**, Hyde RT, Wing AL, *et al*. Physical activity, all-cause mortality, and longevity of college alumni. *N Engl J Med* 1986;**314**:605-13.
- 40 **American College of Sports Medicine**. *ACSM's guidelines for exercise testing and prescription*. 5th ed. Philadelphia: Lippincott, Williams & Wilkins, 2001.
- 41 **Health Canada**. *Handbook for Canada's physical activity guide*. Ottawa, ON.: Health Canada, 2000.
- 42 **US Surgeon General**. *Physical activity and health*. Washington, DC: USPHS Department of Health and Human Services, 1996.
- 43 **American College of Sports Medicine**. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Med Sci Sports Exerc* 1998;**30**:975-91.
- 44 **Armstrong LE**, Maresh CM. The induction and decay of heat acclimatisation in trained athletes. *Sports Med* 1991;**12**:302-12.
- 45 **Shephard RJ**. Immune changes induced by exercise in an adverse environment. *Can J Physiol Pharmacol* 1998;**76**:539-46.
- 46 **Pate RR**, Macera CA. Risks of exercising: musculo-skeletal injuries. In: Bouchard C, Shephard RJ, Stephens T, eds. *Physical activity, fitness and health*. Champaign, IL: Human Kinetics, 1994:1008-18.
- 47 **Shephard RJ**. Exercise and relaxation in health promotion. *Sports Med* 1997;**23**:211-17.
- 48 **Lamb KL**, Brodie DA. The assessment of physical activity by leisure-time physical activity questionnaires. *Sports Med* 1990;**10**:159-80.
- 49 **Godin G**, Shephard RJ. A simple method to assess exercise behavior in the community. *Canadian Journal of Applied Sport Sciences* 1985;**10**:141-6.
- 50 **Haskell WL**, Taylor HL, Wood PD, *et al*. Strenuous physical activity, treadmill exercise test performance and plasma high-density lipoprotein cholesterol. The Lipid Research Clinics Program Prevalence Study. *Circulation* 1980;**62**:53-61.
- 51 **Mundal R**, Erikssen J, Rodahl K. Assessment of physical activity by questionnaire and by personal interview with particular reference to fitness and coronary mortality. *Eur J Appl Physiol* 1987;**56**:245-52.
- 52 **Stephens T**, Craig C. *The well-being of Canadians: the 1988 Campbell's survey*. Ottawa: Canadian Fitness and Lifestyle Research Institute, 1990.
- 53 **Buskirk ER**, Harris D, Mendez J, *et al*. Comparison of two assessments of physical activity and a survey method for caloric intake. *Am J Clin Nutr* 1971;**24**:1119-25.
- 54 **Weiss TW**, Slater CH, Green LW, *et al*. The validity of single-item, self-assessment questions as measures of adult physical activity. *J Clin Epidemiol* 1990;**43**:1123-9.
- 55 **Sallis JF**, Buono MJ, Roby JJ, *et al*. Seven-day recall and other physical activity self-reports in children and adolescents. *Med Sci Sports Exerc* 1993;**25**:99-108.
- 56 **Philippaerts RM**, Westertep KR, Lefevre J. Comparison of two questionnaires with a triaxial accelerometer to assess physical activity patterns. *Int J Sports Med* 2001;**22**:34-9.
- 57 **Caspersen CJ**, Bloemberg PM, Saris WHM, *et al*. The prevalence of selected physical activities and their relation with coronary heart disease risk factors. *Am J Epidemiol* 1991;**133**:1078-91.
- 58 **Kohl HW**, Blair SN, Paffenbarger RS, *et al*. A mail survey of physical activity habits as related to measured physical fitness. *Am J Epidemiol* 1988;**127**:1228-39.
- 59 **Hopkins WG**, Wilson NC, Russell DG. Validation of the physical activity instrument for the life in New Zealand national survey. *Am J Epidemiol* 1991;**133**:73-82.
- 60 **Dan AJ**, Wilbur JE, Hedricks C, *et al*. Lifelong physical activity and older women. *Psychol Women Q* 1990;**14**:531-42.
- 61 **Friedenreich CM**, Courneya KS, Bryant HE. The lifetime total physical activity questionnaire: development and reliability. *Med Sci Sports Exerc* 1998;**30**:266-74.
- 62 **Kriska AM**, Black-Sandler R, Cauley JA, *et al*. The assessment of historical physical activity and its relation to adult bone parameters. *Am J Epidemiol* 1988;**127**:1053-63.
- 63 **Vuillemin A**, Guillemin F, Denis G, *et al*. A computer-assisted assessment of lifetime physical activity: reliability and validity of the QUANTAP software. *Rev Epid Santé Publ* 2000;**48**:157-67.
- 64 **Gretebeck RJ**, Montoye HJ. Variability of some objective measures of physical activity. *Med Sci Sports Exerc* 1992;**24**:1167-72.
- 65 **Weiner JS**, Lourie JA. *Practical human biology*. London: Academic Press, 1981.
- 66 **Uitenbroek DG**. Seasonal variation in leisure time physical activity. *Med Sci Sports Exerc* 1993;**25**:755-60.
- 67 **Baranowski T**. Validity and reliability of self-report measures of physical activity: an information processing perspective. *Res Q* 1988;**59**:314-27.
- 68 **Durante R**, Ainsworth BE. The recall of physical activity using a cognitive model of the question-answering process. *Med Sci Sports Exerc* 1996;**28**:1282-91.
- 69 **Baranowski T**, Sworkin RJ, Cieslik CJ. Validity of children's self report of aerobic activity. *Res Q* 1984;**55**:309-17.
- 70 **Sarkin JA**, Nichols JF, Sallis JF, *et al*. Self-report measures and scoring protocols affect prevalence estimates of meeting physical activity guidelines. *Med Sci Sports Exerc* 2000;**32**:149-56.
- 71 **Montoye HJ**, Block WD, Metzner HL, *et al*. Habitual physical activity and serum lipids: males aged 16-64 in a total community. *Journal of Chronic Diseases* 1976;**29**:697-709.
- 72 **Reiff GG**, Montoye HJ, Remington RD, *et al*. Assessment of physical activity by questionnaire and interview. *J Sports Med Phys Fitness* 1967;**7**:135-42.
- 73 **Wessel JA**, Montoye HJ, Mitchell H. Physical activity assessment by recall record. *Am J Public Health* 1965;**55**:1430-6.
- 74 **Vuillemin A**, Oppert JM, Guillemin F, *et al*. Self-administered questionnaire compared with interview to assess past-year physical activity. *Med Sci Sports Exerc* 2000;**32**:1119-24.
- 75 **Guillemin F**, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures. Literature review and proposal guidelines. *J Clin Epidemiol* 1993;**46**:1417-32.
- 76 **Klesges RC**, Eck LH, Mellon MW, *et al*. The accuracy of self-reports of physical activity. *Med Sci Sports Exerc* 1990;**22**:690-7.
- 77 **Vuillemin A**. Quantification de l'activité physique passée: mesure et rôle pronostique sur l'état de santé (masse osseuse et fonction musculaire) [Quantification of past physical activity and prognostic role with respect to health (bone mass and muscle function)]. PhD, Université Henri Poincaré-Nancy 1, 1998.
- 78 **Sims J**, Smith F, Duffy A, *et al*. The vagaries of self-reports of physical activity: a problem revisited and addressed in a study of exercise promotion in the over 65s in general practice. *Fam Pract* 1999;**16**:152-7.
- 79 **Ching PLYH**, Dietz WH. Reliability and validity of activity measures in preadolescent girls. *Pediatr Exerc Sci* 1995;**7**:389-99.
- 80 **Shephard RJ**. Assessment of physical activity and energy needs. *Am J Clin Nutr* 1989;**50**:1195-200.
- 81 **Bernstein M**, Sloutskis D, Kumanyika S, *et al*. Data-based approach for developing a physical activity frequency questionnaire. *Am J Epidemiol* 1998;**147**:147-54.
- 82 **Washburn RA**. Assessment of physical activity in older adults. *Res Q* 2000;**71**:79-88.
- 83 **Schuit AJ**, Schouten EG, Westertep KR, *et al*. Validity of the Physical Activity Scale for the Elderly (PASE): according to the energy expenditure assessed by the doubly-labeled water method. *J Clin Epidemiol* 1997;**50**:541-6.
- 84 **Sallis JF**, Saelens BE. Assessment of physical activity by self report: status, limitations and future directions. *Res Q* 2000;**71**:1-14.
- 85 **Pennebaker JW**. Psychological bases of symptom reporting and emotional aspects of chemical sensitivity. *Toxicol Ind Health* 1994;**10**:497-511.
- 86 **Thune I**, Lund E. Physical activity and risk of colorectal cancer in men and women. *Br J Cancer* 1996;**73**:1134-40.
- 87 **Borg G**. The perception of physical performance. In: Shephard RJ, ed. *Frontiers of fitness*. Springfield, IL: CC Thomas, 1971:280-94.
- 88 **Goode RC**, Mertens R, Shaiman S, *et al*. Voice, breathing, and the control of exercise intensity. *Adv Exp Med Biol* 1998;**450**:223-9.
- 89 **Tudor-Locke CE**, Myers AM. Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Med* 2001;**31**:91-100.
- 90 **Blair SN**, Haskell WL, Ho P, *et al*. Assessment of habitual physical activity by a seven-day recall in a community survey and controlled experiments. *Am J Epidemiol* 1985;**122**:794-804.
- 91 **Blair SN**, Cheng Y, Holder JS. Is physical activity or physical fitness more important in defining health benefits? *Med Sci Sports Exerc* 2001;**33**:S379-99.
- 92 **Hakim AA**, Curb JD, Petrovich H, *et al*. Effects of walking on coronary heart disease in elderly men: the Honolulu Heart Program. *Circulation* 1999;**100**:9-13.
- 93 **Rakowski W**, Mor V. The association of physical activity with mortality among older adults in the Longitudinal Study of Aging (1984-1988). *J Gerontol* 1992;**47**:M122-9.
- 94 **Wannamethee SG**, Shaper AG. Physical activity in the prevention of cardiovascular disease. An epidemiological perspective. *Sports Med* 2001;**31**:101-14.
- 95 **Durnin JVGA**, Passmore R. *Energy, work and leisure*. London: Heinemann Educational Books, 1967.
- 96 **Parker DL**, Leaf DA, McAfee SR. Validation of a new questionnaire for the assessment of leisure time physical activity. *Ann Sports Med* 1988;**4**:72-81.
- 97 **Lamb KL**. Correlates of self-perceived fitness. *Percept Mot Skills* 1992;**74**:907-14.
- 98 **Lamb KL**, Morris PG. Leisure-time physical activity as a determinant of self-perceived fitness. *Percept Mot Skills* 1993;**76**:1043-7.
- 99 **Blair SN**, Kannel WB, Kohl HW, *et al*. Surrogate measures of physical activity and physical fitness. *Am J Epidemiol* 1989;**129**:1145-56.
- 100 **Heil DP**, Freedson PS, Ahlquist LE, *et al*. Nonexercise regression models to estimate peak oxygen consumption. *Med Sci Sports Exerc* 1995;**27**:599-606.

- 101 **Jackson AS**, Blair SN, Mahar MT, *et al.* Prediction of functional aerobic capacity without exercise testing. *Med Sci Sports Exerc* 1990;**22**:863–70.
- 102 **Lamb KL**, Brodie DA. Leisure-time physical activity as an estimate of physical fitness: a validation study. *J Clin Epidemiol* 1991;**44**:41–52.
- 103 **Kolkhorst FW**, Dolgener FA. Nonexercise model fails to predict aerobic capacity in college students with high VO₂ peak. *Res Q* 1994;**85**:78–83.
- 104 **Whaley MH**, Kaminsky LA, Dwyer GB, *et al.* Failure of predicted VO_{2peak} to discriminate physical fitness in epidemiological studies. *Med Sci Sports Exerc* 1995;**27**:85–91.
- 105 **Shephard RJ**. Normal levels of activity in Canadian city dwellers. *Can Med Assoc J* 1967;**96**:912–14.
- 106 **Goode RC**, Virgin A, Romet T, *et al.* Effects of a short period of physical activity in adolescent boys and girls. *Canadian Journal of Applied Sport Sciences* 1976;**1**:241–50.
- 107 **Welk GJ**, Corbin CB, Dale D. Measurement issues in the assessment of physical activity in children. *Res Q* 2000;**71** [suppl 2]:S59–73.
- 108 **Montoye HJ**, VanHuss WD, Olson HW, *et al.* The longevity and morbidity of college athletes. Lansing, MI: Phi Epsilon Kappa Fraternity, 1957.
- 109 **Paffenbarger RS**, Hyde RT, Wing AL, *et al.* The association of changes in physical activity level and other lifestyle characteristics with mortality among men. *N Engl J Med* 1993;**328**:538–45.
- 110 **Layne JE**, Nelson ME. The effects of progressive resistance training on bone density: a review. *Med Sci Sports Exerc* 1999;**31**:25–30.
- 111 **Vuori I**. Peak bone mass and physical activity: a short review. *Nutr Rev* 1996;**54**:S11–14.
- 112 **Taoli E**, Barone J, Wynder EL. A case-control study on breast cancer and body mass. The American Health Foundation/Division of Epidemiology. *Eur J Cancer* 1995;**31A**:723–8.
- 113 **Lee IM**, Hsieh CC, Paffenbarger RS. Exercise intensity and longevity in men: The Harvard Alumni Health Study. *JAMA* 1995;**273**:1179–84.
- 114 **Washburn RA**, Heath GW, Jackson AW. Reliability and validity issues concerning large-scale surveillance of physical activity. *Res Q* 2000;**71** [suppl 2]:S104–13.
- 115 **Slater CH**, Green LW, Vernon SW, *et al.* Problems in estimating the prevalence of physical activity from national surveys. *Prev Med* 1987;**16**:107–18.
- 116 **Brownson RC**, Jones DA, Pratt M, *et al.* Measuring physical activity with the behavioral risk factor surveillance system. *Med Sci Sports Exerc* 2000;**32**:1913–18.
- 117 **Rennie KL**, Wareham NJ. The validation of physical activity instruments for measuring energy expenditure: problems and pitfalls. *Public Health Nutr* 1998;**1**:265–71.
- 118 **Cohen J**. A coefficient of agreement for nominal scales. *Education and Psychological Measurement* 1960;**20**:37–46.
- 119 **Booth ML**, Okely AD, Chey T, *et al.* The reliability and validity of the physical activity questions in the WHO health behavior in schoolchildren (HBSC) survey: a population study. *Br J Sports Med* 2001;**35**:263–7.
- 120 **Lowther M**, Mutrie N, Loughlan C, *et al.* Development of a Scottish physical activity questionnaire: a tool for use in physical activity interventions. *Br J Sports Med* 1999;**33**:244–9.
- 121 **Bland JM**, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307–10.
- 122 **Weller IMR**, Corey PN. A study of the reliability of the Canada Fitness Survey questionnaire. *Med Sci Sports Exerc* 1998;**30**:1530–6.
- 123 **Folsom AR**, Jacobs DR, Caspersen CJ, *et al.* Test-retest reliability of the Minnesota leisure time physical activity questionnaire. *Journal of Chronic Diseases* 1986;**39**:505–11.
- 124 **Ainsworth BE**, Leon AS, Richardson MT, *et al.* Accuracy of the College Alumnus Physical Activity Questionnaire. *J Clin Epidemiol* 1993;**46**:1403–11.
- 125 **Cauley JA**, LaPorte RE, Kuller LH, *et al.* Comparison of methods to measure physical activity in post-menopausal women. *Am J Clin Nutr* 1987;**45**:14–22.
- 126 **Stewart AL**, Mills KM, King AC, *et al.* CHAMPS Physical activity questionnaire for older adults: outcomes for interventions. *Med Sci Sports Exerc* 2001;**33**:1126–41.
- 127 **Aaron DJ**, Kriska AM, Dearwater SR, *et al.* The epidemiology of leisure physical activity in an adolescent population. *Med Sci Sports Exerc* 1993;**25**:847–53.
- 128 **Slattery ML**, Jacobs DR. Assessment of ability to recall activity of several years ago. *Ann Epidemiol* 1995;**5**:292–6.
- 129 **Kriska A**. Ethnic and cultural issues in assessing physical activity. *Res Q* 2000;**71** [suppl 2]:S47–53.
- 130 **Shephard RJ**. *Men at work*. Springfield, IL: CC Thomas, 1974.
- 131 **Hoyt RW**, Jones TE, Stein TP, *et al.* Doubly labeled water measurement of human energy expenditure during strenuous exercise. *J Appl Physiol* 1991;**71**:16–22.
- 132 **Black AE**, Cole TJ. Within- and between-subject variation in energy expenditure as measured by the doubly-labelled water technique: implications for validating reported dietary energy intake. *Eur J Clin Nutr* 2000;**54**:386–94.
- 133 **Gibney ER**. Energy expenditure in disease: a time to revisit? *Proc Nutr Soc* 2000;**59**:199–207.
- 134 **Philippaerts RM**, Westertker KP, Lefevre J. Doubly labelled water validation of three physical activity questionnaires. *Int J Sports Med* 1999;**20**:284–9.
- 135 **Melanson EL**, Freedson PS. Physical activity assessment: A review of methods. *Crit Rev Food Sci Nutr* 1996;**36**:385–96.
- 136 **Bassett DR**, Ainsworth BE, Swartz AM, *et al.* Validity of four motion sensors in measuring moderate intensity physical activity. *Med Sci Sports Exerc* 2000;**32** [suppl 9]:S471–80.
- 137 **Haskell WL**, Yee MC, Evans A, *et al.* Simultaneous measurement of heart rate and body motion to quantitate physical activity. *Med Sci Sports Exerc* 1993;**25**:109–15.
- 138 **Ainsworth BE**, Jacobs DR, Leon AS. Validity and reliability of self-reported physical activity status: the Lipid Research Clinics questionnaire. *Med Sci Sports Exerc* 1993;**25**:92–8.
- 139 **Richardson MT**, Leon AS, Jacobs DR, *et al.* Ability of the Caltrac accelerometer to assess daily physical activity levels. *J Cardiopulm Rehabil* 1995;**15**:107–13.
- 140 **Voorrips LE**, Ravelli AC, Dongelmans PC, *et al.* A physical activity questionnaire for the elderly. *Med Sci Sports Exerc* 1991;**23**:974–9.
- 141 **Suzuki I**, Kawakami N, Shimizu H. Reliability and validity of a questionnaire for assessment of energy expenditure and physical activity in epidemiological studies. *J Epidemiol* 1998;**8**:152–9.
- 142 **LaPorte RE**, Kuller LH, Kupfer DJ, *et al.* An objective measure of physical activity for epidemiologic research. *Am J Epidemiol* 1979;**109**:158–68.
- 143 **Washburn RA**, Ficker JL. Physical Activity Scale for the Elderly (PASE): the relationship with activity measured by a portable accelerometer. *J Sports Med Phys Fitness* 1999;**39**:336–40.
- 144 **Pols MA**, Peeters PHM, Kemper HCG, *et al.* Repeatability and relative validity of two physical activity questionnaires in elderly women. *Med Sci Sports Exerc* 1996;**28**:1020–5.
- 145 **Steele BG**, Holt L, Belza B, *et al.* Quantitating physical activity in COPD using a triaxial accelerometer. *Chest* 2000;**117**:1359–67.
- 146 **Otis RB**, Brown AS, Womack CJ, *et al.* Relationship between physical activity recall and free-living daily physical activity in older claudicants. *Angiology* 2000;**51**:181–8.
- 147 **Singh PN**, Tonstad S, Abbey DE, *et al.* Validity of selected physical activity questions in white Seventh-day Adventists and non-Adventists. *Med Sci Sports Exerc* 1996;**28**:1026–37.
- 148 **Knapik J**, Bielenda C, Zoltick J, *et al.* Relationships between self-reported physical and physical fitness in active men. *Am J Prev Med* 1993;**9**:203–8.
- 149 **Elosua R**, Garcia M, Aguilar A, *et al.* Validation of the Minnesota leisure time physical activity questionnaire in Spanish women. Investigators of the MARATDON Group. *Med Sci Sports Exerc* 2000;**32**:1431–7.
- 150 **Paffenbarger RS**, Blair SN, Lee IM, *et al.* Measurement of physical activity to assess health effects in free-living populations. *Med Sci Sports Exerc* 1993;**25**:60–70.
- 151 **Young DR**, Jee SH, Appel LJ. A comparison of the Yale Physical Activity Survey with other physical activity measures. *Med Sci Sports Exerc* 2001;**33**:955–61.
- 152 **Siconolfi SF**, Lasater TM, Snow RK, *et al.* Self-reported physical activity compared with maximal oxygen uptake. *Am J Epidemiol* 1985;**122**:101–5.
- 153 **Leon AS**, Jacobs DR, DeBacker G, *et al.* Relationships of physical characteristics and life habits to treadmill exercise capacity. *Am J Epidemiol* 1981;**113**:653–60.
- 154 **Sobolski J**, DeBacker G, Degré S, *et al.* Physical activity, physical fitness and cardiovascular diseases: design of a prospective epidemiological study. *Cardiology* 1981;**67**:38–51.
- 155 **Sobolski J**, Kolesar JJ, Kornitzer MD, *et al.* Physical fitness does not reflect physical activity patterns in middle-aged workers. *Med Sci Sports Exerc* 1988;**20**:6–13.
- 156 **Yasin S**, Alderson MR, Marr JW, *et al.* Assessment of habitual activity apart from occupation. *Br J Prev Soc Med* 1967;**21**:163–9.
- 157 **LaPorte RE**, Black-Sandler R, Cauley JA, *et al.* The assessment of physical activity in older women: an analysis of the inter-relationships and reliability of activity monitoring, activity surveys and caloric intake. *J Gerontol* 1983;**38**:394–7.
- 158 **Johansen KL**, Painter P, Kent-Braun JA, *et al.* Validation of questionnaires to estimate physical activity and functioning in end-stage renal disease. *Kidney Int* 2001;**59**:1121–7.
- 159 **Bairey-Merz CN**, Olson M, McGorray S, *et al.* Physical activity and functional capacity measurement in women: a report from the NHLBI-sponsored WISE study. *J Women's Health* 2000;**9**:769–77.
- 160 **Wolf AM**, Hunter DJ, Colditz GA, *et al.* Reproducibility and validity of a self-administered physical activity questionnaire. *Int J Epidemiol* 1994;**23**:991–9.
- 161 **Aaron DJ**, Kriska AM, Dearwater SR, *et al.* Reproducibility and validity of an epidemiological questionnaire to assess past year physical activity in adolescents. *Am J Epidemiol* 1995;**142**:191–201.
- 162 **Bassett DR**, Cureton AL, Ainsworth BE. Measurement of daily walking distance: questionnaire versus pedometer. *Med Sci Sports Exerc* 2000;**32**:1018–23.
- 163 **Bassett DR**, Vachon JA, Kirkland AO, *et al.* Energy cost of stair climbing and descending on the college alumnus questionnaire. *Med Sci Sports Exerc* 1997;**29**:1250–4.
- 164 **Lee MM**, Whittemore AS, Jung DL. Reliability of recalled physical activity, cigarette smoking and alcohol consumption. *Ann Epidemiol* 1992;**2**:705–14.
- 165 **Simons-Morton DG**, Calfas KJ, Oldenburg B, *et al.* Effects of interventions in health care settings on physical activity or cardiorespiratory fitness. *Am J Prev Med* 1998;**15**:413–30.
- 166 **Young DR**, Haskell WL, Taylor CB, *et al.* Effect of community health education on physical activity knowledge, attitudes and behavior. *Am J Epidemiol* 1996;**144**:264–74.

- 167 **McDermott MM**, Liu K, O'Brien E, *et al.* Measuring physical activity in peripheral arterial disease: a comparison of two physical activity questionnaires with an accelerometer. *Angiology* 2000;**51**:91–100.
- 168 **Tudor-Locke C**. A preliminary study to determine instrument responsiveness to change with a walking program: physical activity logs vs. pedometers. *Res Q* 2001;**72**:288–92.
- 169 **Caspersen CJ**, Merritt RK, Stephens T. International physical activity patterns: a methodological perspective. In: Dishman RK, ed. *Exercise adherence*. Champaign, IL: Human Kinetics, 1994:73–110.
- 170 **Gordis L**. Assuring the quality of questionnaire data in epidemiological research. *Am J Epidemiol* 1979;**109**:21–4.
- 171 **Greenland S**. Dose response and trend analysis in epidemiology: alternatives to categorical analysis. *Epidemiology* 1995;**6**:356–65.
- 172 **Pollock ML**, Gaesser GA, Butcher JD, *et al.* The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness and flexibility in healthy adults. *Med Sci Sports Exerc* 1998;**30**:975–91.
- 173 **Lee I-M**, Paffenbarger RS. Is vigorous physical activity needed to reduce the risk of cardiovascular disease? In: Leon AS, ed. *Physical activity and cardiovascular health: a national consensus*. Champaign, IL: Human Kinetics Publishers, 1997:67–75.
- 174 **Lee I-M**, Paffenbarger RS. Associations of light, moderate and vigorous intensity physical activity with longevity. The Harvard Alumni Health Study. *Am J Epidemiol* 2000;**151**:293–9.
- 175 **Lee I-M**, Skerrett PJ. Physical activity and all-cause mortality: what is the dose-response relationship? *Med Sci Sports Exerc* 2001;**33**(suppl):S459–71.

..... COMMENTARY

Numerous studies on physical activity and health have been published. There are so many tools used to measure physical activity that the problem of validity, reliability, and sensitivity arises. Despite the limitations of questionnaires, they are often used, especially for large populations or studies investigating influence of lifetime physical activity. This paper gives an overview of the limitations and points out the difficulties associated with measurement of physical activity. Different types and patterns of physical activity are detailed, which are used to provide various indicators. It would also be interesting to have an inventory of these indicators and their use depending on the aims of the study. Another pattern that it is important to consider is the regularity of practice during specific periods of life. Bearing in mind the limitations of questionnaire studies, we are able to specify more adapted tools and to interpret more cautiously the results of the studies.

A Vuillemin,

Université Henri Poincaré - Nancy 1, Faculté du Sport, 30 rue du Jardin Botanique, 54600 Villers-les-Nancy, France; Anne.Vuillemin@staps.uhp-nancy.fr



Have your say

eLetters

If you wish to comment on any article published in the *British Journal of Sports Medicine* you can send an eLetter using the eLetters link at the beginning of each article. Your response will be posted on *British Journal of Sports Medicine* online within a few days of receipt (subject to editorial screening).

www.bjsportmed.com