Physical activity is a complex and difficult exposure to assess in epidemiological studies. Improving the accuracy of its assessment is important so that the sub-dimension of this multi-dimensional exposure that is most closely associated with a particular health outcome can be identified. This precision is necessary to develop appropriate interventions that target the specific sub-dimension of interest. Obtaining better information about the strength of association will lead to better estimates of the benefit that may result from successful behaviour change.

For reasons of practicality, most epidemiological studies have relied on questionnaires to assess activity. Many have focused on recreational rather than total activity, probably because it is easier to recall repeated discrete activities that are undertaken for a limited period of time and for which a conscious choice is made prior to engagement. The focus on recreational activity may be appropriate in studies which are occupationally defined or which have samples drawn from restricted socioeconomic groups, because in such studies there will be little variability in...
occupational activity and most of the heterogeneity between individuals will be in recreational activity. Such an argument would not hold, however, in populations with greater variation in work-related physical activity. The focus on recreational activity may, however, represent that element of activity that is most easily modifiable.

Although the totality of physical activity may be the aetologically important exposure for some endpoints, this has rarely been assessed by questionnaires. In their review of physical activity instruments, Montoye et al. noted that few questionnaires had been designed to assess physical activity in work, recreation and domestic life.

This paper describes the development, validation and repeatability testing of a comprehensive instrument designed to measure the different sub-dimensions of physical activity in the Norfolk cohort of the European Prospective Investigation into Cancer (EPIC-Norfolk). As this study is designed to measure associations of lifestyle factors with major chronic diseases in mid-to-late life, including cancer, cardiovascular disease, type 2 diabetes and osteoporosis, the goal was to develop a questionnaire capable of assessing the key sub-dimensions of activity likely to be related to these endpoints; aerobic intensity, overall energy expenditure and load bearing. The resulting EPIC Physical Activity Questionnaire (EPAQ) goes under the name of EPAQ2 to distinguish it from the questionnaire used in the main EPIC study throughout Europe which was a short global assessment of activity derived from a Dutch questionnaire.

Few previous physical activity validation studies have used objective methods as comparisons. The validation instruments selected for this study include a measure of cardio-respiratory fitness and of total energy expenditure as assessed by heart rate monitoring with individual calibration. This method has previously been compared to the gold standard methods of doubly-labelled water and whole body calorimetry, with which it has a high correlation (0.93). Both the assessment of cardio-respiratory fitness and the level and pattern of energy expenditure were repeated three times at 4-monthly intervals throughout the 12-month frame of reference of the questionnaire. The validation study was undertaken in a subgroup of participants in a population-based cohort study in Ely, Cambridgeshire. The participants in this study were of the same age, social class and ethnicity as the Norfolk population for whom the questionnaire was designed. The assessment of reliability took place within the EPIC-Norfolk cohort itself.

Methods

Questionnaire development

The EPAQ2 is a self-completed questionnaire that collects past-year self-reported physical activity behaviours in a disaggregated way such that the information may be re-aggregated according to the dimension of physical activity that is of interest. The questionnaire consists of three sections: activity at home, work and recreation. In each case questions are closed rather than open-ended, to make them easy to complete and to facilitate large-scale data entry. A copy of the questionnaire may be obtained from the EPIC-Norfolk website at http://www.srl.cam.ac.uk/epic/questionnaires/epaq2/. The EPAQ2 asks about activities in the context in which they are undertaken. As an example, other physical activity questionnaires ask people to report how far they walk. This may not be a construct to which most people can easily relate. Instead EPAQ2 asks people about walking in the different sections of their life: in and around home, to work, how much at work, and walking for leisure. The questionnaire went through nine iterations with a series of groups of unselected volunteers in a continuing population-based study. Although the final version of questionnaire has closed questions with ordered categories of continuous variables that the volunteer has to select from, the selection of these categories was determined by the range of responses in early versions where the question was open framed. In the recreation section, which is derived from the Minnesota Leisure Time Activity questionnaire, activities are ordered according to the previously reported frequency in a UK population. Where different intensities of activity are possible within a specific category e.g. running, then the more intense (e.g. competitive running) is asked before the less intense (jogging). The occupational activity is derived from the Modified Tecumseh Occupational Activity questionnaire that has been validated elsewhere.

The validity of the questionnaire was assessed in a subgroup of 173 participants in the existing population-based Ely study and its reliability was measured in a subgroup of 399 participants within the EPIC-Norfolk study itself.

Validity study

Selection of study participants

The volunteers in the validity study were all participants in the Ely cohort, a continuing population-based study in Ely, Cambridgeshire, the detailed design of which has been described previously. In the second phase of the Ely Study, 790 individuals completed measurements of anthropometry, cardiorespiratory fitness and 4-day energy expenditure by heart rate monitoring. A random sub-set of 200 individuals from this cohort were asked to re-attend a further three times over the following year when all tests were repeated. Overall, 173 of the participants completed all measurements and therefore had four measures of cardio-respiratory fitness and four measures of 4-day energy expenditure by heart rate monitoring (Figure 1) completed across one year. At the final visit, participants completed the EPAQ2 questionnaire that asks about physical activity in the past year. Ethical permission for the study was granted by the Cambridge Local Research Ethics Committee.

Measurement protocol

On each of the four visits, height and weight were measured in light clothing and body fat percentage was obtained using a

![Time intervals for the assessment of objective measures of physical activity and assessment of past year physical activity by the questionnaire (EPAQ2)](image-url)
standard impedance technique (Bodystat, Isle of Man). Body circumferences were measured using a metal tape. The waist circumference was measured at the mid-point between the lower costal margin and the level of the anterior superior iliac crest. Hip circumference was measured at the level of the greater trochanter. The protocol for undertaking the individual calibration between heart rate and energy expenditure has been reported previously.6,16 The oxygen consumption-heart rate relationship was assessed at rest with the subject lying prone and then seated, using an oxygen analyser calibrated daily using 100% nitrogen and fresh air as standard gases. To provide the slope and the intercept of the line relating energy expenditure to heart rate, each participant cycled at 50 revolutions per minute and the workload was progressively increased from 0 W, through 37.5 W, 75 W and 125 W in stages each lasting 5 minutes. At each workload three separate readings were made of heart rate, minute volume and expired air oxygen concentration. The 125 W level was only undertaken if the heart rate had not reached 120 beats per minute by the end of the 5 minutes at 75 W. The oxygen concentration in the expired air and minute volume data were used to calculate oxygen consumption after correction for standard temperature and pressure. Energy expenditure (kJ/min) was calculated at each time point as oxygen consumption (ml/min) × 20.35.17 Mean resting energy expenditure was taken as the average of the lying and sitting values. Flex heart rate, the empirical point at which the distinction between rest and exercise is made, was calculated as the mean of the highest resting pulse rate and the lowest on exercise. Finally the slope and intercept of the least squares regression line of the exercise points were calculated. Maximum oxygen uptake (VO₂max) was measured from the linear regression as predicted oxygen consumption at maximal heart rate (220-age) and is expressed in the results per unit body weight. The volunteers wore the heart rate monitor (Polar Electro, Finland) continuously during the waking hours over the following 4 days. Heart rate readings were directly downloaded into a computer via a serial interface and the individual calibration data were used to predict minute energy expenditure for each person. The energy expenditure data were summed over the day to create an estimate of daytime energy expenditure (Daytime EE). Finally the minute-by-minute energy expenditure data was analysed to calculate the proportion of time when the physical activity ratio (PAR), the ratio of minute energy expenditure to minute basal metabolic rate (BMR), was >5.18 This was used in the analysis as an indicator of the pattern of physical activity. Basal metabolic rate was calculated from prediction equations.18 A cut-off of five was used because this energy intensity ratio discriminates moderate and highly intense activities from those that are light.19 For each individual the means of body mass index (BMI), percentage body fat, waist-hip ratio (WHR), daytime energy expenditure and VO₂max on the four occasions were calculated and are used in the analysis as the measures of the usual level of obesity and its regional distribution, energy expenditure and fitness, respectively.

**Repeatability study**

**Selection of study participants**

The repeatability of this questionnaire was assessed in participants in the EPIC-Norfolk study. At baseline, 1993–1997, EPIC-Norfolk recruited 25,633 people aged 45–74 years. In 1997 a follow-up study was commenced which finally resulted in repeat visits for 15,786 members of this cohort. By May 1999, 4853 people had attended for the second health check and had full data entry. They are used in the paper as the representative cohort from which the repeatability sub-sample was selected. At the follow-up visit participants provided a blood and urine sample and completed a detailed health and lifestyle questionnaire and the EPAQ2 questionnaire. Beginning in January 1999, 544 volunteers were asked at the end of their visit whether they would complete a second physical activity questionnaire that would be posted to them in 3 months time. This interval was selected in order to have a high degree of overlap between the periods covered without the interval being too short such that participants simply recalled their first response when completing the questionnaire on the second occasion.20 In all, 498 individuals agreed to complete the second questionnaire which was sent to them at three months. In all, 399 individuals returned completed questionnaires.

**Statistical analysis**

Estimates of energy expenditure at home, work and during recreation from the questionnaire were calculated by multiplying participation (h/week) by the metabolic cost of each activity, expressed in metabolic equivalents (MET) obtained from published tables.18,19 The sum in each category is given in units of MET·h per week. A MET is the ratio of the energy cost of a given activity to resting metabolic rate. In addition we calculated time spent viewing television and videos (h/week) and reported time participating in more vigorous recreational activities i.e. those with a MET score >5.

1-t test comparisons between men and women for mean anthropometric, energy expenditure and fitness measurements were undertaken. All the questionnaire-derived physical activity variables, apart from hours spent watching television, were skewed and comparison between men and women was performed on log-transformed data. The relationship between the questionnaire-derived physical activity variables and the objective measures, daytime energy expenditure, VO₂max per kg and proportion of time with PAR > 5 was described using correlation coefficients with and without adjustment for age and sex. The correlation between repeated assessments of the questionnaire-derived physical activity variables was described using Pearson correlation. In addition, participants were classified into quartiles of the activity indices. If the raw activity score was equal to zero for more than 25% of the distribution then the baseline category constituted all subjects whose score was zero with the remaining positive scores being split into tertiles. If the number of participants with a zero score was less than 25% of the distribution then the distribution was divided into quartiles. Cohen’s weighted kappa, a measure of agreement, was calculated according to the four resulting categories between the baseline and the repeat questionnaire.20

**Results**

**Validity study**

The 173 participants in the validation study were middle-aged men and women with BMI and proportions that are typical of those seen in adults of this age in EPIC-Norfolk (Table 1). Average daytime energy expenditure as assessed by heart rate
monitoring was higher in men than women. Men also had a significantly higher maximum oxygen uptake. Comparison of the mean VO$_{2\text{max}}$ data to those seen in the Allied Dunbar National Fitness Survey suggests that the validation population is not any fitter than the general population.\textsuperscript{13} External comparison of the questionnaire-derived data with other populations is difficult, as limited national data are available.

Mean reported television viewing was not significantly different between men and women. Reported energy expenditure at home was significantly greater in women than in men, whereas the reverse was true for activity at work. Total recreational activity was higher in men than in women, but there was no significant difference in time spent in vigorous recreational activity. Overall 58% of men and 54% of women participated in no vigorous activity.

Pearson's partial correlation coefficients adjusted for age and sex between the objective measures of physical activity and variables derived from the questionnaire are given in Table 2. The questionnaire indices of energy expenditure in recreation and at work showed modest positive correlation with daytime energy expenditure from heart rate monitoring. However, correlation between reported activity at home and daytime energy expenditure was negatively correlated. Reported activity at home was negatively correlated with reported activity at work and recreation (partial correlation coefficients adjusted for age and sex, –0.23 and –0.22, respectively). Thus an overall measure of physical activity summing domestic, occupational and recreational activity was less correlated with daytime energy expenditure than either occupational or recreational activity separately (data not shown). Therefore in constructing a summary index, the self-reported physical activity index, only occupational and recreational activity were summed. This self-reported physical activity index was significantly positively correlated with mean daytime energy expenditure after adjustment for age and sex ($r = 0.28$, $P < 0.001$). The unadjusted ($r = 0.44$, $P < 0.001$) and age-adjusted data ($r = 0.45$, $P < 0.001$) were more strongly correlated. In sex-specific analyses adjusted for age, the correlation in men was $0.30$ ($P < 0.01$) and $0.23$ in women ($P < 0.05$).

Time spent watching television was negatively associated with our objective measure of cardio-respiratory fitness (VO$_{2\text{max}}$ per kg).
and the coefficient was unchanged either in the unadjusted analysis or after adjustment for age and sex ($r = -0.24$ and $r = -0.19$, respectively). Reported energy expenditure in recreational activity and time spent participating in vigorous activities were both positively correlated with $\text{VO}_{2\text{max}}$ per kg and also the proportion of time spent with PAR $\geq 5$, our objective measure of the pattern of vigorous activity. The correlation of these questionnaire parameters of vigorous activity was less strong with mean daytime energy expenditure, as would have been predicted.

**Repeatability study**

Table 3 shows a comparison of the questionnaire-derived physical activity variables in the larger EPIC cohort compared to the sub-set on whom a second EPAQ2 was collected. There were no statistically significant differences between the main EPIC-Norfolk cohort and the repeatability sub-sample. In men, the correlation coefficients for the various indices between the repeat questionnaires were $\geq 0.68$ for all indices except work activity, which had a correlation coefficient of 0.57 (Table 4). In women, the correlation coefficients were $> 0.60$ for all indices except work activity and amount of time spent participating in vigorous sporting activities which were 0.37 and 0.41, respectively. All of the correlations were statistically significant at the 0.05 level.

**Discussion**

The questionnaire described in this paper was designed to assess the energy expenditure and fitness enhancing elements of physical activity, a complex and difficult exposure to measure in epidemiological studies. The results suggest that the repeatability of the indices computed from the questionnaire is high but that validity compared to repeated objective estimates of energy expenditure measured throughout the frame of reference of the questionnaire is modest. The 3-month interval between repeats of the questionnaire in the repeatability study, chosen to reduce the probability that individuals recall their previous response, might be considered relatively long. However, the effect of a lengthy re-test interval would be to diminish rather than exaggerate the correlation between repeats. Thus the high values of repeatability demonstrated represent, at worse, an underestimate of the true repeatability. The measurement of validity of physical activity questionnaires is difficult, but our study design allows for assessment with objective and independent comparison instruments repeated throughout the frame of reference of the questionnaire. The study group selected for the validation study were chosen at random from a continuing population-based cohort study. As such they tend to be unselected with regard to physical activity behaviour. By contrast in the closest comparable study, the SAFE study, the participants

**Table 3** Repeatability study: descriptive characteristics and questionnaire-derived physical activity variables in the EPIC cohort ($n = 4853$) and repeat sample ($n = 399$)

<table>
<thead>
<tr>
<th></th>
<th>Men Total cohort ($n = 2126$)</th>
<th>Men Repeat sample ($n = 187$)</th>
<th>Women Total cohort ($n = 2727$)</th>
<th>Women Repeat sample ($n = 212$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)$^a$</td>
<td>64.6 (8.4)</td>
<td>65.0 (8.2)</td>
<td>63.2 (8.5)</td>
<td>63.8 (8.4)</td>
</tr>
<tr>
<td>Body mass index (kg/m$^2$)$^a$</td>
<td>26.9 (3.6)</td>
<td>26.9 (3.3)</td>
<td>26.8 (4.9)</td>
<td>26.9 (4.0)</td>
</tr>
<tr>
<td>TV viewing (h/wk)$^a$</td>
<td>21.6 (9.8)</td>
<td>24.6 (10.5)</td>
<td>22.5 (9.9)</td>
<td>24.3 (11.4)</td>
</tr>
<tr>
<td>Activity at home (MET.h/wk)$^b$</td>
<td>18.1 (9.1,30.4)</td>
<td>20.6 (12.9,32.0)</td>
<td>58.5 (41.2,78.2)</td>
<td>55.2 (41.2,74.0)</td>
</tr>
<tr>
<td>Activity at work (MET.h/wk)$^b$</td>
<td>0.0 (0.0,70.4)</td>
<td>0.0 (0.0,15.7)</td>
<td>0.0 (0.0,15.8)</td>
<td>0.0 (0.0,16.8)</td>
</tr>
<tr>
<td>Recreational activity (MET.h/wk)$^b$</td>
<td>27.8 (14.7,48.0)</td>
<td>28.4 (14.9,47.0)</td>
<td>17.6 (7.4,31.8)</td>
<td>16.7 (8.4,28.8)</td>
</tr>
<tr>
<td>Vigorous activity (h/wk)$^b$</td>
<td>0.0 (0.0,0.8)</td>
<td>0.0 (0.0,0.5)</td>
<td>0.0 (0.0,0.7)</td>
<td>0.0 (0.0,0.7)</td>
</tr>
<tr>
<td>Self-reported physical activity index (MET.h/wk)$^b$</td>
<td>54.9 (25.0,108.1)</td>
<td>43.0 (20.1,86.2)</td>
<td>26.7 (11.0,57.4)</td>
<td>24.7 (10.8,49.0)</td>
</tr>
</tbody>
</table>

Values shown are $^a$ mean and standard deviation and $^b$ median and interquartile range.

**Table 4** Repeatability study: Pearson correlation coefficients and Cohen weighted kappa statistics for 3-month repeat of questionnaire-derived variables of physical activity ($n = 399$)

<table>
<thead>
<tr>
<th></th>
<th>Men Total cohort ($n = 2126$)</th>
<th>Men Repeat sample ($n = 187$)</th>
<th>Women Total cohort ($n = 2727$)</th>
<th>Women Repeat sample ($n = 212$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TV time (h/wk)</td>
<td>$\rho$ 0.75</td>
<td>$\kappa$ 0.74</td>
<td>$\rho$ 0.75</td>
<td>$\kappa$ 0.74</td>
</tr>
<tr>
<td>Activity at home (MET.h/wk)</td>
<td>$\rho$ 0.77</td>
<td>$\kappa$ 0.74</td>
<td>$\rho$ 0.77</td>
<td>$\kappa$ 0.74</td>
</tr>
<tr>
<td>Activity at work (MET.h/wk)</td>
<td>$\rho$ 0.61</td>
<td>$\kappa$ 0.62</td>
<td>$\rho$ 0.61</td>
<td>$\kappa$ 0.62</td>
</tr>
<tr>
<td>Recreational activity (MET.h/wk)</td>
<td>$\rho$ 0.69</td>
<td>$\kappa$ 0.64</td>
<td>$\rho$ 0.69</td>
<td>$\kappa$ 0.64</td>
</tr>
<tr>
<td>Vigorous activity (h/wk)</td>
<td>$\rho$ 0.54</td>
<td>$\kappa$ 0.55</td>
<td>$\rho$ 0.54</td>
<td>$\kappa$ 0.55</td>
</tr>
<tr>
<td>Self-reported physical activity index (MET.h/wk)</td>
<td>$\rho$ 0.75</td>
<td>$\kappa$ 0.70</td>
<td>$\rho$ 0.75</td>
<td>$\kappa$ 0.70</td>
</tr>
</tbody>
</table>
were paid volunteers, mostly graduates employed in administrative or professional positions. The participants in our validation study were on average 5 years younger than the population for whom the questionnaire was designed. However, this is unlikely to result in a biased assessment of validity as the physical activity patterns of the two groups are likely to be similar, containing a mixture of working age and retired people. The relatively low validity may be a reflection of the difficulty in capturing the between-individual differences in energy expenditure in a simple questionnaire. One of the obvious limitations is that a single estimate of the energy cost of an activity taken from a published compendium is applied to all individuals.

This does not allow for any between-individual variation in energy expenditure for a given activity and so cannot take account of differences in intensity (e.g. through people playing sport at different competitive levels), or through variations in mechanical and metabolic efficiency. The use of the compendium values of energy expenditure is also limited by the validity of the estimates for each activity as they are derived from studies on small groups of selected volunteers. However, the published estimates of energy costs of activities are the only such data that are available.

The validity and reliability results are similar to those previously reported for other questionnaires. In a review of available questionnaires, Kriska and colleagues noted how most, but not all, previously published questionnaires were supported by reliability studies and where this was measured, it was almost always high. By contrast, validity was less frequently reported and where it was, it was typically low. One particular problem for the validation of physical activity questionnaires is the choice of appropriate comparison instrument. Many physical activity validation studies have used other forms of subjective questionnaire or diary as the validation method. Although this strategy tends to produce higher correlations, the possibility of correlated error is substantial as both the questionnaire under scrutiny and the validation instrument are of the same fundamental type and are subject to the same forms of bias. It would be preferable therefore to select an objective non-questionnaire based method as the validation instrument, ideally one with a high correlation to the true exposure of interest. As physical activity behaviour is related to several objective non-questionnaire based methods as the validation instrument, ideally one with a high correlation to the true exposure of interest. As physical activity behaviour is related to several objective non-questionnaire based methods as the validation instrument, ideally one with a high correlation to the true exposure of interest. As physical activity behaviour is related to several objective non-questionnaire based methods as the validation instrument, ideally one with a high correlation to the true exposure of interest. As physical activity behaviour is related to several objective non-questionnaire based methods as the validation instrument, ideally one with a high correlation to the true exposure of interest.

In summary, the repeatability and validity studies on EPAQ2 suggest that it is a useful tool for assessing physical activity in large epidemiological studies such as EPIC-Norfolk. As with any physical activity questionnaire, the inference about its validity and reliability is limited to the population in whom it was tested and its use is restricted to the purpose for which it was intended. The logical development and construction of a questionnaire plays a key role in determining its validity. It is therefore appropriate to describe development alongside validation.

Acknowledgements

The Ely Study was funded by the Medical Research Council and Anglia and Oxford Regional Health Authority. We are grateful to H Shannasy, S Curran, P Murgatroyd, and Drs M Hennings and AM Prentice for their help with the fieldwork for the Ely study and to Suzie Oakes, Robert Luben, Joanna Camus and the EPIC-Norfolk field team for their assistance. NJW is an MRC Clinician Scientist Fellow.

A copy of the EPAQ2 questionnaire can be downloaded from the EPIC-Norfolk website at http://www.srl.cam.ac.uk/epic/questionnaires/epaq2/
KEY MESSAGES

- Physical activity is an important but difficult exposure to assess in epidemiological studies.
- The choice of comparison method and selection of study population are key determinants of the inferences that can be drawn from validation studies of physical activity questionnaires.
- The reliability of the European Prospective Investigation into Cancer Study-Norfolk cohort (EPIC-Norfolk) physical activity questionnaire (EPAQ2) has been demonstrated in a sub-study within EPIC itself.
- The validity of EPAQ2 has been demonstrated in a study in a comparable population using repeated objective assessment of energy expenditure and fitness throughout the frame of reference of the questionnaire.

References