

Estimation of Reproducibility and Relative Validity of the Questions Included in the EPIC Physical Activity Questionnaire

MARGREET A POLS,* PETRA H M PEETERS,* MARGA C OCKÉ,** NADIA SLIMANI,†
H BAS BUENO-DE-MESQUITA** AND HUBERTINE J A COLLETTE*

Pols M A (Department of Epidemiology, University of Utrecht, PO Box 80035, 3508 TA Utrecht, the Netherlands), Peeters P H M, Ocké M C, Slimani N, Bueno-de-Mequita H B and Collette H J A. Estimation of reproducibility and relative validity of the questions included in the EPIC Physical Activity Questionnaire. *International Journal of Epidemiology* 1997; **26** (Suppl. 1): S181–S189.

Background. The EPIC core questionnaire on lifestyle contains a number of questions on physical activity designed to rank subjects according to level of physical activity (short PA questionnaire). These questions are based on a more extensive questionnaire designed to measure absolute total energy expenditure (extensive PA questionnaire), that was validated in a pilot study preceding EPIC. Reproducibility and relative validity of the short PA questionnaire were estimated by selecting, from the pilot study data, the answers to a number of questions from the extensive questionnaire that resembled those actually included in the short version.

Methods. The population of the pilot study consisted of 126 men and women aged between 20 and 70 years. Reproducibility was estimated by administering the extensive questionnaire three times: at baseline, and after 5 and 11 months. In order to determine the relative validity of the extensive questionnaire, a 3-day activity diary, repeated four times, was used as the reference method.

Results. Over the study period (13 months), mean absolute energy expenditure, estimated from the questions included in the short questionnaire, was fairly constant in men but not in women. **Reproducibility:** Spearman correlation coefficients ranged from 0.47 to 0.89 in men, and from 0.49 to 0.81 in women. **Relative validity:** Spearman correlation coefficients between the short questionnaire and the diary were between 0.32 and 0.81 for men, and between 0.28 and 0.72 for women.

Conclusions. The questions selected for the short questionnaire are not suitable for estimating energy expenditure at an absolute level. Reproducibility and relative validity of the ranking of subjects seemed satisfactory and comparable to the extensive questionnaire. The results imply that the short questionnaire is suitable for ranking subjects in the EPIC study.

Keywords: EPIC, relative validity, reproducibility, physical activity

During the past few decades a growing interest in the relationship between physical activity and the occurrence of chronic disease has developed. Lack of physical activity has been found to be associated with a variety of diseases. The relationship most frequently described is an inverse one with cardiovascular disease. Physical activity is believed to have a beneficial effect on non-insulin-dependent diabetes mellitus and osteoporosis.¹

Although most studies find an elevated risk for inactive people for colon cancer, results for breast and prostate cancer are not very consistent.² It is likely that risk estimates are weakened by random errors in measurement. But even if relative risks are rather low, the attributable risk could be very large because of the large number of people with a sedentary lifestyle. Therefore, it was decided to collect data on physical activity in the European Prospective Investigation into Cancer and Nutrition (EPIC). This study, which was started in 1993, involves seven European countries (France, Germany, Greece, Italy, the Netherlands, Spain, and the United Kingdom), representing a total cohort of about 350 000 subjects.³ Because of this international aspect, physical activity patterns will show relatively large variations.

* Department of Epidemiology, University of Utrecht, PO Box 80035, 3508 TA Utrecht, the Netherlands.

** Department of Chronic Disease and Environmental Epidemiology, National Institute for Public Health and the Environment, PO Box 1, 3720 BA Bilthoven, the Netherlands.

† Unit of Analytical Epidemiology, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon, Cedex 08 France.

The methods commonly used to assess physical activity or energy expenditure, especially in clinical research, are not always appropriate for large epidemiological studies (for example: calorimetry, movement sensors). In a large-scale epidemiological study the assessment method should be repeatable, valid, user-friendly, and inexpensive. The method most frequently used is the questionnaire. Questionnaires can differ in time period of reference, nature and detail of activities, mode of data collection (interviewer or self-administered), and method of computing energy expenditure, or an activity index.⁴ All methods suitable for large population studies are prone to measurement error. It is therefore important to assess the validity of each method in the population for which it will be used.

A physical activity questionnaire was developed for use in EPIC, the aim being to estimate usual daily total energy expenditure in the course of the past year. This questionnaire is further referred to as the 'pre-EPIC questionnaire' (preliminary or extensive). This was validated in the Dutch pilot study performed during the period October 1991–October 1992. However, although the results of the validation study were relatively good,⁵ problems of feasibility and costs led to the development of a smaller number of physical activity questions, based on the extensive questionnaire, for inclusion in the EPIC core (or short) questionnaire. These questions will be referred to here as the short questionnaire. The aim of this study is to estimate reproducibility and relative validity of the actual EPIC PA questions, by selecting from the pilot study data the answers to a number of questions from the extensive questionnaire that resembled the short questionnaire.

METHODS

The Dutch pilot study, which was called BALANS, was conducted in the period October 1991–October 1992. In this study, both the extensive questionnaire and the Dutch food frequency questionnaire were tested. Results on reproducibility and relative validity of the food frequency questionnaire are reported in companion papers.^{6,7} To assess the reproducibility of the extensive questionnaire, it was administered three times (November 1991, April 1992, October 1992). Relative validity was determined by comparing the results of the questionnaire with those of a 3-day activity diary repeated four times: in April, May, June and July, 1992.

Subjects

An age-stratified sampling frame was defined consisting of 260 women (age 50–70 years) who were invited to take part in a breast cancer screening programme

TABLE 1 *Characteristics of the study population*

	Men (n = 64)		Women (n = 62)	
	Mean	SD ^a	Mean	SD
Age (years)	41.1	11.0	48.8	14.8
Height (m)	1.79	0.05	1.65	0.07
Weight (kg)	81.3	9.4	67.3	9.4
QI (kg/m ²) ^b	25.4	2.8	24.8	3.5

^a SD = standard deviation.

^b QI = Quetelet index.

in Utrecht and a population of 700 men and women (age 20–59 years) who participated in a monitoring programme on risk factors for chronic diseases in the cities of Amsterdam, Doetinchem and Maastricht. Subjects were recruited from these two ongoing projects, because the Dutch EPIC study is conducted within these projects. Out of 960 subjects invited to participate, 240 (25%) responded positively, 288 (30%) refused to participate, and 432 (45%) did not respond. Of the 240 subjects who responded positively, people who either did not speak Dutch or were not available for 13 consecutive months were excluded. Eventually, 134 subjects were selected, equally distributed over the four cities, both sexes, and over 20-year age groups. Since the intended population size was defined at about 120 subjects (as advised by Willett for dietary validation studies⁸), no attempt was made to obtain co-operation from the non-responders. All participants signed an informed consent form.

There were six subjects whose activity patterns changed markedly during the study year: five of them stopped working during this period while one restarted formal training, which implied a change in activity pattern. Two subjects dropped out of the study in October 1991, before the first extensive questionnaire was administered. These eight subjects were excluded from all analyses. Table 1 presents the characteristics of the 126 subjects included in the analyses. The average age of the female participants was higher than that of the male, due to the difference in age structure between the two source populations.

During the course of the study another 10 subjects dropped out. This, together with a varying number of subjects who missed one or more questions, resulted in a different number of subjects for the three consecutive times the extensive questionnaire was administered. The number of subjects for every measurement is therefore mentioned in the tables.

TABLE 2 Comparison of the actual EPIC physical activity questions (short questionnaire) with corresponding questions from the pre-EPIC (extensive) questionnaire

Short questionnaire	Extensive questionnaire
1. Choose what best corresponds to your present occupation: sedentary occupation, standing occupation, manual work, heavy manual work	What is your profession (open-ended)? Estimation of hours/week spent on light, moderate, heavy, and very heavy work
2. In a typical week during the past year, how many hours per week did you spend on:	
a. Walking, including going to work, shopping, leisure time	Average duration one-way trip to work on foot How many days a week do you work? How much time do you spend walking for pleasure? Average duration one-way trip to work per bicycle How many days a week do you work? How much time do you spend working in the garden? How much time do you spend on do-it-yourself jobs? Do you regularly participate in sport? Name of regular sport, frequency, duration Name of seasonal sport, frequency, duration Number of hours per day/week/month spent on all households tasks
b. Cycling, including going to work and leisure time	
c. Gardening	
d. Do-it-yourself activities at home	
e. Physical exercise	
f. Housework	
3. Sweating/faster heartbeat	
4. In a typical week during the past year, how many flights of stairs did you climb per day?	How many flights of stairs do you climb per day?

e.g. Walking, 2.a. in the short questionnaire, is estimated by C.3. (job frequency) times B.2. (walking to or from work) plus F.1. (walking for pleasure) in the extensive questionnaire.

Short questionnaire

The short questionnaire is presented in the Appendix; it is self-administered. Revision of the extensive questionnaire was considered necessary because of its length. In addition, the objectives of the questionnaire were changed: instead of measuring total absolute individual energy expenditure, the present aim is to select and measure the most important indicators of variation in energy expenditure in order to obtain a relatively good ranking of subjects according to their level of physical activity. Resting and sleeping were expected to account for little within- and between-person variability, and were therefore excluded. Since job activity is fairly constant, subjects are asked to classify it on a scale from 1 to 4. For other activities (walking, cycling, gardening, do-it-yourself-activities, sports, and housework) subjects are asked to estimate the time spent on these activities in a normal week, and whether these activities lead to perspiring (or faster heartbeat). A separate question was added about the number of flights of stairs subjects climb per week.

The actual questions on the short questionnaire were not tested in the pilot study. In order to get an impression whether the ranking of subjects using these questions is worse compared to the full pre-EPIC questionnaire, the answers to those pre-EPIC questions resembling the actual EPIC PA questions were selected.

Reproducibility and relative validity of those selected EPIC PA questions were estimated, and compared to the results of the full pre-EPIC questionnaire.

Extensive questionnaire

The questionnaire validated in the pilot study was developed to measure in a standardized way the total individual energy expenditure. The subjects filled in the average number of hours spent during the past year on different activities of the following types: resting, transport to or from work, work, household activities, sports, and other activities. The questionnaire was self-administered and was checked by a trained dietician (who also checked the food frequency questionnaires) at the study centre in the presence of and in consultation with the subject.

Energy expenditure was computed by multiplying the amount of time spent on specific activities by values for energy expenditure per activity per unit time according to James and Schofield.⁹ The questions from the extensive questionnaire that were used to estimate the reproducibility and relative validity of the short questionnaire are listed in Table 2.

Diary

The diaries for assessment of relative validity were based on the method of Bouchard.¹⁰ In the diaries for

the BALANS study, subjects registered their activities in periods of 5 minutes, assigning to every period a letter indicating the activity (e.g. R for resting, S for standing). In this study, subjects were instructed to use predefined moments of the day (breakfast, coffee break, lunch, etc.) to write down the activities in the previous hours. If an activity could not be assigned to one of the coded categories (resting, sitting, standing, walking, cycling, driving a car, house-keeping, dressing), either a T or an X was entered: T stands for training or playing sports and X is an open-ended activity report. If either T or X was used, the subjects were asked to note at the bottom of the page which activity was meant by T or X. The activities registered under X were mainly gardening or do-it-yourself-activities. On average, recorded days were evenly distributed over all 7 days of the week.

Daily energy expenditure was computed using the energy cost for each category as estimated from reference values according to James and Schofield.⁹ Activities described under either T or X were classified afterwards into four groups according to energy cost. For each group a mean energy expenditure was estimated and used to compute energy expenditure per 24 hours.

Of the 126 participants included, 115 filled in activity diaries; 103 subjects completed all 12 days, seven subjects 9 days, three subjects 6 days, and two subjects 3 days. Daily energy expenditure was computed as the mean expenditure of the recorded days.

Data Analysis

Statistical analyses were performed using SPSS/PC+ 4.0, and SPSS for Windows 6.1. Means and standard deviations were computed for the energy expenditure estimated from the short questionnaire and from the diaries for those categories of activity that were alike in both methods. Difference-against-mean plots¹¹ were drawn for total energy expenditure from the short questionnaire estimated from measurements from the first and the third extensive questionnaire for both men and women (the plot for women is presented).

Since not all variables were normally distributed, for assessment of reproducibility Spearman correlation coefficients (5 and 11 months) were computed between energy expenditure calculated from the short questionnaire, selected from the repeated administrations of the extensive questionnaire. Estimation of reproducibility for the job question was not possible, because in the extensive questionnaire job activity was not coded 1 to 4 as it is in the short questionnaire. Instead, subjects were asked to estimate the number of hours spent on sedentary work, standing work, heavy work, or very heavy work. It would mean too much interpretation to

assign a code from 1 to 4 to those answers, and then to assess the reproducibility of these codings.

For estimation of relative validity Spearman's correlation coefficients between energy expenditure derived from the short questionnaire and from the average of all four activity diaries were calculated. Relative validity was computed for the first time and the third time the extensive questionnaire was filled in. The first questionnaire was chosen to avoid the study effect, because in the EPIC study subjects also complete the questionnaire only once. However, because the questionnaire was designed to measure physical activity in the preceding year, the correlation may be underestimated. The third questionnaire referred to the year in which the diaries were recorded and the dietary recalls were collected. Therefore, results for the third questionnaire are also presented. This correlation, on the other hand, may be an overestimation because the subjects were more aware after being frequently asked about their activity patterns. Thus, the true correlation may lie somewhere in between.

Relative validity of both the EPIC job question and climbing stairs were not assessed, because these activities were not coded separately in the diaries (job activity was coded as walking, sitting, etc.). This was also the case for gardening and do-it-yourself activities. But since most activities recorded in the open-ended category (X) in the diaries were described as gardening or do-it-yourself, the correlation between the sum of these two questions and the diary open-ended category was calculated.

RESULTS

There was little difference in total energy expenditure calculated from the three extensive questionnaires (Table 3), indicating that on the group level absolute energy expenditure did not change during the study period. Total energy expenditure estimated from the short questionnaire was constant for men but decreased for women.

Reproducibility

Figure 1 shows a difference-against-mean plot for energy expenditure estimated from the short questionnaire selected from the first and third administration of the extensive questionnaire, for women. The mean difference was 734 kJ. The so-called limits of agreement (mean \pm 2SD) are -3530 and 4998 kJ. This means that about 95% of the outcomes of the third measurement falls between approximately 3500 kJ between and 5000 kJ above the first measurement. For men the

TABLE 3 Means and standard deviations (SD) for the extensive questionnaire, the short questionnaire, and activity diaries, for men and women

		Men		Women	
		Mean ± SD	n	Mean ± SD	n
Extensive questionnaire	November	9455 ± 2282 kJ (2251 ± 543 kcal)	63	8491 ± 1984 kJ (2005 ± 472 kcal)	59
	April	9661 ± 2412 kJ (2300 ± 574 kcal)	63	8416 ± 1613 kJ (2004 ± 384 kcal)	53
	October	9489 ± 3018 kJ (2259 ± 719 kcal)	58	8593 ± 1600 kJ (2033 ± 381 kcal)	55
Short questionnaire	November	3429 ± 2549 kJ (816 ± 607 kcal)	63	4145 ± 2710 kJ (987 ± 645 kcal)	60
	April	3393 ± 2456 kJ (808 ± 585 kcal)	63	3647 ± 2480 kJ (868 ± 590 kcal)	60
	October	3380 ± 2895 kJ (805 ± 689 kcal)	63	3434 ± 2250 kJ (818 ± 536 kcal)	62
Activity diaries (12 days)		12 204 ± 1478 kJ (2906 ± 352 kcal)	60	11 960 ± 1085 kJ (2848 ± 258 kcal)	55

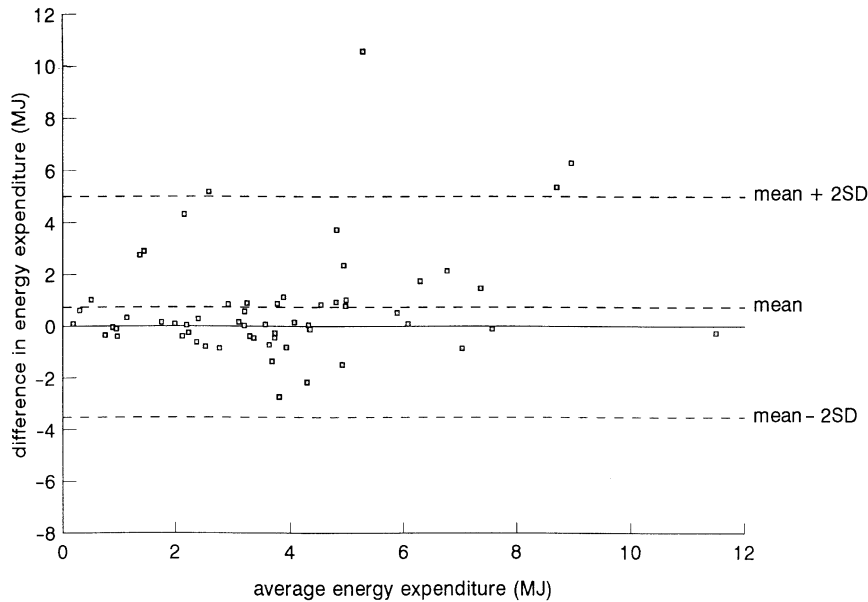


FIGURE 1 Difference in absolute energy expenditure against average energy expenditure estimated from the short questionnaire (first and third administration), for women

mean difference was 50 kJ, with limits of agreement of -3879 kJ and 3977 kJ.

Spearman correlation coefficients (Table 4) for men ranged from 0.58 to 0.89 (5 months) and from 0.47 to 0.83 (11 months). For women, correlations were between 0.59 and 0.81 (5 months) and between 0.49 and 0.72 (11 months). On average, 5 months reproducibility was better than 11 months.

Relative Validity

Total energy expenditure estimated from the short questionnaire accounted for approximately 30% of total

energy expenditure computed from the diary. Also for separate activity categories estimates of energy expenditure from the short questionnaire were lower than those calculated from the diary, except for housework, which was higher (data not shown).

The Spearman correlations of the short questionnaire (selected from the first administration of the extensive questionnaire) with the activity diaries were between 0.32 and 0.81 for men and between 0.28 and 0.72 for women (Table 5). Correlations between total energy expenditure estimated from the short questionnaire (first administration) and energy expenditure from the diary

TABLE 4 Reproducibility of the short questionnaire: 5 and 11 months Spearman correlation coefficients for energy expenditure, men and women^a

	Men		Women	
	5 months (n)	11 months (n)	5 months (n)	11 months (n)
Walking	0.58 (63)	0.57 (63)	0.64 (59)	0.57 (61)
Cycling	0.73 (63)	0.47 (63)	0.77 (58)	0.72 (60)
Gardening	0.77 (64)	0.71 (64)	0.81 (62)	0.67 (62)
DIY ^b	0.81 (64)	0.83 (64)	0.80 (62)	0.54 (62)
Sports	0.73 (64)	0.51 (64)	0.61 (62)	0.54 (62)
Housework	0.89 (64)	0.77 (64)	0.64 (62)	0.72 (62)
Stairs	0.69 (64)	0.66 (64)	0.59 (62)	0.49 (62)
Total energy	0.86 (63)	0.75 (63)	0.63 (58)	0.68 (60)

^a 95% confidence intervals for $r = 0.60$ (0.41–0.74), $r = 0.70$ (0.54–0.81), $r = 0.80$ (0.69–0.88), $r = 0.90$ (0.84–0.94) for $n = 60$.

^b DIY = do-it-yourself activities.

were 0.43 for men and 0.51 for women. For the third administration results for total energy were slightly better. For both sexes, housework correlated best with diary energy expenditure, and in both groups the correlations for walking were quite low. There was no consistent difference in correlations with the diaries between the first and the third administrations.

DISCUSSION

This pilot study demanded a lot from the participants. Since no attempt was made to obtain co-operation from the non-responders, it is possible that the 25% of invited subjects who took part in the study is a selected group in that they are probably more aware of their health. Data on non-responders are limited. They did not differ in age distribution from the responders (χ^2 test). For part of the invited population more characteristics are available, but it is not known what variables one should use to identify a health-oriented lifestyle. If the population is a selected group, this could have led to overestimation of reproducibility and relative validity.¹²

The short questionnaire itself was not tested in this pilot study, but the extensive questionnaire was. In these analyses the answers to those questions from the extensive questionnaire that resembled the short questionnaires were used to estimate the reproducibility and relative validity of the latter. The aim was to determine whether the ranking of subjects using the short questionnaire was sufficiently reproducible and valid.

TABLE 5 Relative validity: Spearman correlation coefficients between the short questionnaire and diary energy expenditure, first and third administration^a

	Men		Women	
	First (n)	Third (n)	First (n)	Third (n)
Walking	0.32 (59)	0.26 (59)	0.33 (54)	0.33 (55)
Cycling	0.46 (59)	0.48 (59)	0.48 (53)	0.49 (55)
Gardening + DIY ^b	0.43 (60)	0.50 (60)	0.47 (55)	0.35 (55)
Sports	0.55 (60)	0.49 (60)	0.28 (55)	0.40 (55)
Housework	0.81 (60)	0.72 (60)	0.72 (55)	0.76 (55)
Total energy	0.43 (59)	0.45 (59)	0.51 (53)	0.56 (55)

^a 95% confidence intervals for $r = 0.30$ (0.04–0.52), $r = 0.40$ (0.15–0.60), $r = 0.50$ (0.27–0.68), $r = 0.60$ (0.40–0.75), $r = 0.70$ (0.53–0.81), $r = 0.80$, (0.68–0.88) for $n = 55$.

^b DIY = do-it-yourself activities.

For energy expenditure assessment the physical activity diary is not a real gold standard. The value of the method depends on the accuracy with which subjects register their activities. In addition, for practical reasons it was not possible to have the subjects fill in the diaries in all seasons. The error this might have caused in the estimation of usual energy expenditure is probably small, since results from the repeated extensive questionnaires suggested that there was no important seasonal influence on energy expenditure. In spite of these limitations, the diary was chosen as a reference method to estimate relative validity. It was considered to be a method with an error which is not correlated with the error of the questionnaire, because time frame and method of recording differed for the two procedures. However, some correlated error could be caused by the fact that the values of energy expenditure per activity for the two methods were taken from the same literature source, and by subject-specific under- or over-reporting in both methods. Another reason for choosing the activity diary was the possibility of identifying separate categories of activity. Apart from being costly, more accurate methods such as either doubly labelled water or activity monitoring devices would only yield total energy expenditure.

In the literature, there are several ways of obtaining reference values for energy expenditure per activity. Some tables list energy expenditure per kg body weight per unit of time. This would mean that a 100 kg person expends twice as much energy during walking than a 50 kg person, which is probably not true. Since in our study population mean body weight was rather high,

this would result in an overestimation of energy expenditure.¹³ Tables that provide reference values for men and women separately assume a different standard body weight for the two sexes, for example 70 kg for men and 60 kg for women. Usually these weights are lower than the body weights of the study population. This would lead to underestimation of energy expenditure. Energy expenditure can also be expressed in METs, multiples of resting metabolic rate.¹⁴ Nevertheless, it is not clear whether the relationship between energy expenditure and resting metabolic rate for a particular activity is the same for all individuals. The development of tables providing sex, age, and weight-specific reference values for energy expenditure per activity would be very useful for studies like ours. In this study it was decided to use the same values for energy expenditure per activity for men and women.⁸ The literature values were based on data from a number of studies. If most of these studies were carried out among men, this will have led to an overestimation of energy expenditure for women, because on average body weight of women is lower than that of men.

Energy expenditure computed from the diaries was higher than that computed from the extensive questionnaire. This is mainly caused by the fact that most subjects reported only 15–20 hours in the extensive questionnaire, while the diaries were designed to cover the full 24 hours.

Reproducibility

The difference-against-mean plot shows that at a group level reproducibility for absolute energy expenditure estimated from the short questionnaire was poor. The reproducibility of ranking of individuals according to energy expenditure was moderate to good. Spearman correlations were comparable to those calculated from the extensive questionnaire (0.58–0.97).

The correlations are consistent with those of other studies. Baecke *et al.*¹⁵ assessed reproducibility of a questionnaire in a population of men and women aged 19–31 years. Test-retest correlations after 3 months were 0.88, 0.81, and 0.74 for the work, sports and leisure time index, respectively. Voorrips *et al.*¹⁶ developed a questionnaire that was based on Baecke's, and assessed reproducibility after 20 days in a population of 29 subjects aged 63–80 years. Spearman's correlation coefficient was 0.89. Lakka and Salonen¹⁷ computed intraclass correlations for several methods administered at baseline and after one year in 51 men aged 54 years. Correlations were 0.58 for a 12-month history and 0.69 for a habitual occupational activity interview. Washburn *et al.*¹⁸ developed a Physical Activity Scale for the Elderly (PASE) and found a 3–7 week test-retest

correlation coefficient of 0.75 in a population of 254 elderly subjects.

In this study, the time span between measurements was considerably longer compared to other studies, which makes it less likely that the subjects remembered has they answered the previous time. However, the questions refer to different time periods (preceding year) and it is possible that individual activity patterns have changed, leading to an underestimation of reproducibility. If there had been a seasonal influence on the way subjects filled in the questionnaire, reproducibility at 11 months would have been better than at 5 months, which was not the case.

Relative Validity

Absolute values of energy expenditure estimated for questions from the short questionnaire differed substantially from the diary categories. In particular, walking was reported much more frequently in the diaries than in the questionnaire. This can be explained by the fact that e.g. walking during work is reported as a job activity in the questionnaire, while it is reported as walking in the diary. For the other categories differences were smaller, but still substantial. However, the short questionnaire was not designed to estimate absolute energy expenditure, but to rank subjects.

Spearman correlations between total energy expenditure estimated from the short questionnaire and from the diary were moderate. Correlations between specific activities and corresponding categories of the activity diaries showed a large variation, ranging for men from 0.26 (walking) to 0.81 (housework), and for women from 0.27 (sports) to 0.76 (housework). In both sex groups the estimated time spent on walking showed poor relative validity, probably caused by the same factor as the difference in absolute energy expenditure. The ranking according to time spent on housework showed the highest relative validity. Apparently people know quite well how much time they dedicate to household activities. The category of do-it-yourself and gardening was not the same in the two methods. For the diaries, the open-ended (X-) category included all activities which did not belong in the other categories. Therefore this category comprises more than do-it-yourself activities and gardening. This probably decreased the correlation.

There was no indication that the results of the November 1991 questionnaire represented an underestimation and those of the October 1992 questionnaire an overestimation of relative validity. Therefore synchronization of time periods of reference for test and reference methods may not be of great importance.

Comparing results with those of other studies is difficult, because of the large differences in methods

and time frames. Although in this study most correlations are poor to moderate, the results of this study are considerably better than those of other studies. Only Voorrips *et al.*¹⁶ found Spearman correlation coefficients of 0.78 for a questionnaire with a repeated physical activity recall and 0.72 with a pedometer in a population of 31 subjects aged 63–80 years. Cauley *et al.*¹⁷ described correlations between several physical activity assessment methods, most of which were below 0.3. Correlations of Washburn's PASE with health status and physiological measures such as grip strength, resting heart rate, age, and subjective health status were all below 0.5. Still the PASE is described as a 'valid instrument for the assessment of physical activity in epidemiologic studies of older people'.¹⁸ Thus, except for Voorrips' study, most studies to validate physical activity questionnaires found rather poor correlation coefficients.

In conclusion, the results on reproducibility and relative validity of the short questionnaire (EPIC) should be regarded with some caution. First, the questions used are not exactly the same as those selected from the extensive (pre EPIC) questionnaire, although most of the activities covered are the same. Second, the pilot study population may have been a selected group of health-conscious subjects. This may have had a positive influence on reproducibility and relative validity estimates. The short questionnaire could not be used to estimate absolute energy expenditure, neither per category, nor total energy expenditure. Values for absolute energy expenditure not only showed poor relative validity, but also poor reproducibility. Reproducibility and relative validity of the ranking were satisfactory and well within the range of values found in other studies. Correlations for the short questionnaire were of the same magnitude as those of the extensive questionnaire. These results suggest that the simplified version of the extensive questionnaire is suitable for ranking subjects in the EPIC study. However, this study was performed in a Dutch population, and results should not be generalized to the whole EPIC cohort. Therefore it is still important to conduct an international validation study of the short (actual EPIC PA) questionnaire.

ACKNOWLEDGEMENTS

The authors thank the 'Stichting Preventicon voor de Vroege Opsporing van Borstkanker in Midden-Nederland' for their co-operation. They also thank the interviewers for data collection, and Gatske Obermann-de Boer and Jet Smit for their support during the

planning and data collection phase of the study. This study was supported financially by the Europe Against Cancer Programme of the Commission of European Communities (WHO/AEP/90/05).

REFERENCES

- Blair S N, Kohl H W, Gordon N F. How much physical activity is good for health? *Ann Rev Public Health* 1992; **13**: 99–126.
- Sternfeld B. Cancer and the protective effect of physical activity: the epidemiological evidence. *Med Sci Sports Exerc* 1992; **24**: 1195–1209.
- Riboli E. Nutrition and cancer: Background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann Oncol* 1992; **3**: 783–91.
- LaPorte R E, Montoye H J, Caspersen C J. Assessment of physical activity in epidemiologic research: problems and prospects. *Public Health Rep* 1985; **100**: 131–46.
- Pols M A, Peeters P H M, Ocké M C *et al.* Relative validity and repeatability of a new questionnaire on physical activity. *Prev Med* (In Press), 1997.
- Ocké M C, Bueno-de-Mesquita H B, Goddijn H E *et al.* The Dutch EPIC food frequency questionnaire. I. Description of the questionnaire and relative validity and reproducibility for food groups. *Int J Epidemiol* 1997; **26** (Suppl. 1): S37–S48.
- Ocké M C, Bueno-de-Mesquita H B, Pols M A, Smit H A, van Staveren W A, Kromhout D. The Dutch EPIC food frequency questionnaire. II. Relative validity and reproducibility for nutrients. *Int J Epidemiol* 1997; **26** (Suppl. 1): S49–S58.
- Willett W. *Nutritional epidemiology*. New York, Oxford: Oxford University Press, 1990.
- James W P T, Schofield E C. Energy cost of physical activity classified in alphabetical order (appendix 4.2). In: *Human Energy Requirements, a Manual for Planners and Nutritionists*. Oxford, New York, Tokyo: Oxford University Press, 1990.
- Bouchard C, Tremblay A, Leblanc C, Lortie G, Savard R, Thériault G. A method to assess energy expenditure in children and adults. *Am J Clin Nutr* 1983; **37**: 461–67.
- Bland J M, Altman D G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **I**: 307–310.
- Goldbohm R A, van den Brandt P A, Brants H A M, van 't Veer P, Al M, Sturmans F, Hermus R J J. Validation of a dietary questionnaire used in a large-scale prospective cohort study on diet and cancer. *Eur J Clin Nutr* 1994; **48**: 253–65.
- Sallis J F, Haskell W L, Wood P D *et al.* Physical activity assessment methodology in the five-city project. *Am J Epidemiol* 1985; **121**: 91–106.
- Ainsworth B E, Haskell W L, Leon A S *et al.* Compendium of physical activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc* 1993; **25**: 71–80.
- Baecke J A H, Burema J, Frijters J E R. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutr* 1982; **36**: 936–42.
- Voorrips L E, Ravelli A C J, Dongelmans P C A, Deurenberg P, van Staveren W A. A physical activity questionnaire for the elderly. *Med Sci Sports Exerc* 1991; **23**: 974–79.

¹⁷ Lakka T A, Salonen J T. Intra-person variability of various physical activity assessments in the Kuopio Ischaemic Heart Disease Risk Factor Study. *Int J Epidemiol* 1992; **21**: 467–72.

¹⁸ Washburn R A, Smith K W, Jette A M, Janney C A. The physical activity scale for the elderly (PASE): development and evaluation. *J Clin Epidemiol* 1993; **46**: 153–62.

¹⁹ Cauley J A, LaPorte R E, Black Sandler R, Schramm M M, Kriska A M. Comparison of methods to measure physical activity in postmenopausal women. *Am J Clin Nutr* 1987; **45**: 14–22.

APPENDIX

Short Questionnaire

1. Work

We would like to know the type and amount of physical activity involved in your work. Please check what best corresponds with your present occupation from the following four possibilities:

- Sedentary occupation

You spend most of your time sitting (such as in an office) _____

or

- Standing occupation

You spend most of your time standing or walking. However, your work does not require intense physical effort (e.g. shop assistant, hairdresser, guard, etc.) _____

or

- Manual work

This involves some physical effort including handling of heavy objects and use of tools (e.g. plumber, electrician, carpenter, etc.) _____

or

- Heavy manual work

This implies very vigorous physical activity including handling of very heavy objects (e.g. docker, miner, bricklayer, construction worker, etc.) _____

2. In a typical week during the past year, how many hours did you spend per week on each of the following activities:

- a) Walking, including walking to work, shopping, and during leisure time

In summer ___ hours per week

In winter ___ hours per week

- b) Cycling, including cycling to work and during leisure time

In summer ___ hours per week

In winter ___ hours per week

- c) Gardening

In summer ___ hours per week

In winter ___ hours per week

- d) Do-it-yourself activities at home

_____ hours per week

- e) Physical exercise such as fitness, aerobics, swimming, jogging, tennis, etc.

In summer ___ hours per week

In winter ___ hours per week

- f) Housework, such as cleaning, washing, cooking, child care, etc.

_____ hours per week

3. In a typical week during the past year, did you engage in any of these activities vigorously enough to cause sweating or faster heartbeat?

No ___ Yes ___

If yes, for how many hours per week in total did you perform vigorous activity?

_____ hours per week

4. In a typical week during the past year, how many flights of stairs did you climb per day?

_____ floors per day