

Physical activity monitoring in Europe. The European Physical Activity Surveillance System (EUPASS) approach and indicator testing

A Rütten^{1,*}, A Vuillemin², WTM Ooijendijk³, F Schena⁴, M Sjöström⁵, T Stahl⁶, Y Vanden Auweele⁷, J Welshman⁸ and H Ziemainz¹

¹University of Erlangen–Nuremberg, Institute of Sport Science, Gebbertstrasse 123, D-91056 Erlangen, Germany:

²University Henri Poincaré–Nancy 1, Nancy, France: ³TNO Prevention and Health, Leiden, The Netherlands:

⁴Centro Interuniversitario Europeo per le Scienze Sportive ECUS, Rovereto, Italy: ⁵Karolinska Institute, Stockholm, Sweden: ⁶Jyväskylän Yliopisto, Jyväskylä, Finland: ⁷Katholiek Universiteit Leuven, Leuven, Belgium: ⁸Lancaster University, Lancaster, UK

Submitted 9 September 2002: Accepted 28 November 2002

Abstract

Objectives: The main objective of this paper is to describe the approach and specific findings of the European Physical Activity Surveillance System (EUPASS) research project. In particular, the analysis presented aims at testing the reliability, comparability and predictive power of different sets of physical activity (PA) indicators.

Design: First, a panel study based on computer-aided telephone interview (CATI) was designed to report PA data of a representative, selected group of about 100 persons per country at three points in time. Second, a CATI time series survey was carried out with the goal of realising about 100 interviews per month over six consecutive months.

Setting: The project was carried out in eight European countries to support the development of the European Union's (EU) Health Monitoring Programme.

Subjects: Random population samples (subjects aged 18 years and older) were drawn from each participating country.

Results: While many PA indicators used in EU countries to date as well as the psychosocial and environmental measures tested in the present study had acceptable to good reliability coefficients, the test–retest reliability scores of the International Physical Activity Questionnaire (IPAQ) version tested (the short (last 7 days) telephone interview IPAQ; IPAQ-S7T) were rather low. The comparability between extant national PA items and the IPAQ-S7T was low for all countries. The strongest predictors of perceived health were the psychosocial and environmental PA indicators.

Conclusions: According to the results of the present study, more research is needed to further investigate and improve the quality of the IPAQ. In addition, the specific predictive power of the tested psychosocial and environmental PA indicators on perceived health should be of particular interest for designing health surveillance activities in the future.

Keywords
Physical activity
Monitoring
Indicator testing
Public health

Two major developments in international public health have encouraged the present project on physical activity (PA) monitoring in Europe. First, there is increasing support from public health authorities in promoting PA to improve the health of the population. The benefits of PA are related to its broad impact on traditional public health indicators such as mortality and morbidity, as well as on psychosocial well-being and quality of life^{1–4}. Moreover, the promotion of PA is considered an efficient intervention strategy as it can affect large proportions of the population that are currently at health risk because of a sedentary

lifestyle^{3,5}. Second, there is increasing interest in promoting international health surveillance activities. For example, the Centers for Disease Control and Prevention in the USA and the Finnish Public Health Institute have recently organised a series of conferences on the issue of global surveillance⁶.

At present, several international research groups are working on the elaboration of valid, international and comparable instruments to monitor PA. One large international co-operation operates globally on the development and testing of an International Physical

*Corresponding author. Email alfred.ruetten@sport.uni-erlangen.de

Activity Questionnaire (IPAQ). Another group focuses on the harmonisation of PA measurement within the European context of Health Interview Surveys (EUROHIS). The present approach to explore a European Physical Activity Surveillance System (EUPASS) concentrates only on member states of the European Union (EU).

EUPASS approach

The specificity of the EUPASS project relates mainly to the development of a European health monitoring system as promoted by the Health Monitoring Programme of the European Commission. Within this context, the general aims of EUPASS were to develop and test a surveillance system* for PA as a major behavioural determinant of health by: (1) providing a valid and cross-nationally applicable list of core indicators and optional indicators for health-enhancing PA; (2) testing selected PA indicators by employing different survey methodologies; and (3) investigating implementation structures of health monitoring in the EU.

To investigate the comparability of PA indicators existing in EU countries and at the same time improve indicator definitions to be used in an EU indicator set, indicator test surveys of existing and new indicators were prepared and conducted in the countries participating in the project. A co-ordinated questionnaire was developed that was compiled in such a way as to maximise the comparability of PA indicators between EU countries and the EU indicator set. It was used in a combined panel and time series design which included a test of feasibility of indicators for employment in different surveillance methodologies.

The EUPASS network encompassed: (1) the project group built by public health research institutions from eight EU member states (Belgium, Finland, France, Germany, Italy, The Netherlands, Spain, United Kingdom); (2) counterparts within national surveillance institutes of participating countries; and (3) co-operation partners of EU member states not directly included in the project group.

In addition, the EUPASS network has established co-operation with other relevant research activities. For example, close contact has been established with both the IPAQ and EUROHIS groups.

This paper focuses on the primary tasks of the EUPASS project as defined within the Health Monitoring Programme, i.e.

1. testing the quality and reliability of (a) PA indicators used in surveillance systems in the member states to date, (b) new, comparable indicators of PA behaviour (IPAQ) and (c) new comparable sets of psychosocial and environmental determinants of PA;
2. testing the comparability of existing and new indicators; and
3. testing the predictive power of the different sets of PA indicators with regard to health status.

For further methodological issues and results of the project, see Rütten *et al.*⁸ and Rzewnicki *et al.*⁹

Design of the indicator test study

For indicator and survey method testing, three surveys were conducted in each of the EUPASS countries (see Table 1). First, a panel study based on computer-aided telephone interview (CATI) was designed to report PA data of a representative, selected group of about 100 persons in each country at three points in time (T1–T3). Data from T1 and T2 were especially used for reliability testing. Second, a CATI time series survey was carried out over six consecutive months with the goal of realising about 100 interviews per month (i.e. a total of ~600) per country. These data have been used to investigate the quality and comparability of the national indicators used to date, by comparing them with the IPAQ indicators as an internationally tested frame of reference for PA measurement. The data also provided an empirical basis to test the predictive power of different sets of indicators (national indicators vs. IPAQ indicators vs. psychosocial and environmental indicators). Third, a mail survey ($n = 100$) was conducted in each country to control for effects of different survey methods (telephone vs. mail).

Design of the EUPASS questionnaire

A list of relevant indicators for assessing PA and its determinants was selected from the material gathered in the context of an inventory of PA indicators used in EU countries to date and the co-operations between EUPASS and other work groups on the assessment of PA and its determinants. On this basis, a questionnaire was compiled and used in all three surveys (panel, continuous, mail). The EUPASS questionnaire contained four sections including the following groups of indicators:

- indicators of PA behaviour from relevant national health monitoring systems (section A of the questionnaire);
- indicators of PA behaviour developed by the IPAQ group (section B of the questionnaire);
- indicators of environmental, social and individual determinants of PA behaviour (section C of the questionnaire); and
- sociodemographic information (section D of the questionnaire).

*According to the current discussion on global health monitoring, the term surveillance refers particularly to: 'the creation of a data system for changing the public health'⁷. Thus, surveillance can be described as a complex organised effort to: (1) continuously collect data, e.g. monitor long-term changes in behaviour risk factors; (2) analyse these data; and (3) feed back results of analysis to potential users, e.g. public health policy-makers.

Table 1 Design of the study

CATI panel survey (repeated measures)	T1 June 2000 <i>n</i> = 100	T2 1–3 weeks after T1 <i>n</i> = 100	T3 October/November 2000 <i>n</i> = 100			
CATI time series (‘continuous’) survey	S1 June 2000 <i>n</i> = 100	S2 July 2000 <i>n</i> = 100	S3 August 2000 <i>n</i> = 100	S4 September 2000 <i>n</i> = 100	S5 October 2000 <i>n</i> = 100	S6 November 2000 <i>n</i> = 100
Mail method (control survey)	October/November 2000 <i>n</i> = 100					

Section A was different for each participating country as it only included PA indicators used in the main health behaviour surveys for that country. For example, in Finland, the four questions from the National Survey on Health Behaviour among Finnish Adults formed section A of the questionnaire; in Italy the nine questions from the National Survey on Aspects of Everyday Life were used, and so on.

Sections B, C and D of the EUPASS questionnaire used the same indicators in all countries.

In section B, in collaboration with the IPAQ group, indicators from the International Physical Activity Questionnaires were included. (For the history and development of the IPAQ, see <http://www.ipaq.ki.se>.) From extant IPAQ versions, the *short telephone version* was chosen for the indicator test in EUPASS for feasibility reasons. Also, it was decided to use the *last 7 days* – rather than the usual week – version. The questions in this version (IPAQ-S7T) concern the frequency and duration of (1) vigorous and (2) moderate PA, (3) walking and (4) sitting. The answers to these single indicators are summed up to produce an overall indicator of PA-related energy expenditure (MET). Furthermore, in this section of the EUPASS questionnaire, a question (four items) asking directly for PA in the context of job, transportation, housework/house maintenance/caring for family and recreation/sports/leisure was added to the IPAQ short version.

Section C of the EUPASS questionnaire included some psychosocial and environmental determinants of PA that had been tested for their predictive power on PA and health in earlier studies. These three questions made use of (1) a self-efficacy scale (three items) based on the work of Sallis and colleagues¹⁰ and De Bourdeaudhuij and co-workers^{11,12}, (2) a social support scale (five items) and (3) a supportive environment scale (three items). Both of the latter two were based on items tested in a European study on health policy and health behaviour.^{13–15}

Section D contained questions mainly related to sociodemographics. The items selected included sex, age, years of education, household income, occupational status, height, weight and perceived health.

For cultural adaptation and translation of questions and items, the procedure proposed by the IPAQ group was agreed upon. To further prepare the indicator test, national versions of the questionnaire were

constructed (including any necessary national adaptation of instruments) and survey infrastructures were set up (staff, sampling procedures, hardware) (for details, see Rzewnicki *et al.*⁹).

Methods

Random population samples were drawn in each of the eight participating countries. Inclusion criteria were resident of a private household within the country and having reached the age of 18 years (for details, see Rzewnicki *et al.*⁹).

The following statistical analyses were applied. First, after testing the distribution of the data from panel surveys T1 and T2, a non-parametric measure (Spearman’s rank correlation) was used to examine the reliability of the different PA indicators. As a further step in the reliability analysis, the test–retest coefficients from the IPAQ-S7T were compared with the test–retest coefficients for the national indicators used in the different countries, as well as with the psychosocial and environmental determinants indicators.

Second, to investigate the comparability of each of the individual indicators, for each country all of the items from the national indicators and the IPAQ-S7T were correlated.

Third, a hierarchical regression analysis was performed to test the predictive power of the different sets of PA indicators with regard to perceived health status.

Results

Table 2 summarises the results of the fieldwork in the different countries. It reveals major differences in overall response rates for telephone and mail surveys between countries, as well as in specific response rates for the telephone versus the mail survey. For example, Finland reported the highest response rates for all three types of survey (panel 51.6%, continuous 54.5%, mail 58.3%), while the lowest response rates for all surveys were reported for the UK (panel 14.5%, continuous 25.5%, mail 18.6%). In Germany, response rates for the continuous telephone survey were comparably high (50.5%) while mail responses were very low (19.1%). In contrast, France did much better on the mail survey (52.4%) than on the continuous telephone survey (29.1%).

Table 2 Sample description

		Panel				Continuous	Mail
		T1	T2	T3	All		
Belgium	Net sample	622	200	102	622	1577	588
	Realised sample	202	102	79	79	611	206
	Response rate (%)	32.5	51.0	77.5	12.7	38.7	35.0
Finland	Net sample	217	151	127	217	1107	230
	Realised sample	151	127	112	112	603	134
	Response rate (%)	69.6	84.1	88.2	51.6	54.5	58.3
France	Net sample	482	140	91	482	2060	250
	Realised sample	140	91	67	67	599	131
	Response rate (%)	29.0	65.0	73.6	13.9	29.1	52.4
Germany	Net sample	951	382	202	951	1293	350
	Realised sample	389	223	145	145	653	67
	Response rate (%)	40.9	58.4	71.8	15.2	50.5	19.1
Italy	Net sample	608	219	121	608	1892	500
	Realised sample	219	121	91	91	600	148
	Response rate (%)	36.0	55.3	75.2	14.9	31.7	29.6
The Netherlands	Net sample	324	124	95	324	1400	426
	Realised sample	124	95	76	76	606	108
	Response rate (%)	38.3	76.6	80.0	23.5	43.3	25.4
Spain	Net sample	276	158	128	276	1284	300
	Realised sample	158	128	100	100	600	22
	Response rate (%)	57.2	81.0	78.1	36.2	46.7	7.3
United Kingdom	Net sample	546	148	120	546	2838	377
	Realised sample	148	120	79	79	723	70
	Response rate (%)	27.1	81.1	65.8	14.5	25.5	18.6
All nations	Net sample	4026	1522	986	4026	13451	3021
	Realised sample	1531	1007	749	749	4995	886
	Response rate (%)	38.0	66.2	75.9	18.6	37.1	29.3

Reliability of IPAQ-S7T, national and other indicators

Table 3 reports test–retest correlation coefficients for the IPAQ-S7T for the eight participating EU countries. In general, most item correlation coefficients ranged between 0.3 and 0.5, which appears to be rather low for a reliability test. Only the question related to the duration (sum of minutes) of sitting during weekdays provided slightly better results (most coefficients were 0.6 or 0.7). The coefficients for the overall indicator of PA (energy

expenditure score, MET) varied from 0.2 for France to about 0.6 for Spain.

Test–retest coefficients for the national indicators turned out to be similar to (e.g. 0.3 to 0.6 for Germany, 0.3 to 0.7 for the UK) or even better (e.g. 0.5 to 0.9 for Finland, 0.5 to 0.8 for Italy) than those of the IPAQ-S7T. The coefficients for the psychosocial and environmental indicators were also slightly better for all nations, ranging from 0.5 to 0.6 for the self-efficacy scale and the social support scale and from 0.6 to 0.7 for the supportive environment scale.

Table 3 Test–retest reliability of IPAQ-S7T (Spearman's correlation coefficients)

	Belgium	Finland	France	Germany	Italy	The Netherlands	Spain	United Kingdom	All nations
<i>n</i> variables (IPAQ)	100	127	91	223	98	86	128	98	951
Vigorous PA (days)	0.553	0.477	0.278	0.508	0.414	0.344	0.540	0.469	0.494
Vigorous PA (total min)	0.442	0.590	0.359	0.536	0.530	0.413	0.616	0.345	0.509
Moderate PA (days)	0.365	0.283	0.181	0.430	0.208	0.402	0.381	0.254	0.364
Moderate PA (total min)	0.385	0.553	0.352	0.536	0.221	0.338	0.322	0.431	0.389
Walking (days)	0.310	0.550	0.358	0.540	0.471	0.292	0.372	0.495	0.468
Walking (total min)	0.703	0.440	0.504	0.328	0.408	0.297	0.721	0.310	0.461
Walking (intensity)	0.399	0.339	0.453	0.223	0.274	0.422	0.679	0.560	0.441
Sitting (weekdays) (total min)	0.521	0.701	0.422	0.642	0.726	0.633	0.618	0.552	0.623
Sitting (weekend) (total min)	0.338	0.640	0.370	0.407	0.333	0.454	0.431	0.435	0.461
Sum MET PA*	0.531	0.405	0.294	0.388	0.135	0.341	0.576	0.499	0.446
Sum MET sitting†	0.418	0.582	0.417	0.523	0.567	0.497	0.504	0.536	0.527
Sum MET total‡	0.561	0.423	0.225	0.293	0.297	0.376	0.563	0.400	0.419

IPAQ – International Physical Activity Questionnaire; IPAQ-S7T – short (last 7 days) telephone interview IPAQ; PA – physical activity.

* MET – energy expenditure score (1 MET = 1 kcal kg⁻¹ h⁻¹); Sum MET PA – sum of METs (vigorous, moderate and walking in the last 7 days).

† Sum MET sitting – sum of METs (sitting weekdays and weekend).

‡ Sum MET total – sum of METs (Sum MET PA + Sum MET sitting).

Comparability of IPAQ-S7T and national indicators

The comparability of PA indicators used in participating countries to date is very low⁸. Generally, the various EU countries have used different concepts of PA, focused on different dimensions, and used different scales and reference periods in their national surveys.

As a major result of the present analysis, the national instruments showed little comparability with the IPAQ instrument as well. For example, by reporting correlation coefficients only of at least 0.3, in Belgium one item (PA that was intense enough to make you sweat) correlated (0.3) with one item of the IPAQ-S7T (days of vigorous activity per week). In Finland also, one item out of four questions (how demanding is job physically) correlated (0.3, 0.4) with two items of the IPAQ-S7T (duration of moderate activity and of sitting during weekdays). In Italy, correlation of the national indicators and the IPAQ-S7T indicators produced no coefficients above 0.3 at all. Slightly better results were obtained for the UK (three old items correlated with three IPAQ-S7T items), Germany (four correlations above 0.3) and The Netherlands (five correlations above 0.3). However, there was only one item with a correlation between the old and new indicators that was above 0.5, i.e. a question about the duration of sitting from Monday to Friday in Germany, which correlated (0.6) with the sitting question item (duration weekdays) in the IPAQ-S7T.

Predictive power of different indicators

The EUPASS project also tested a group of selected new indicators that are especially related to psychosocial and environmental determinants of health. For investigation of the predictive power of the different sets of indicators used in the study (i.e. (1) national indicators, (2) IPAQ indicators and (3) psychosocial and environmental indicators), hierarchical regression analyses were conducted for each country with subjective health status as the dependent variable. This variable was self-rated by respondents, who selected one of the response categories 'very good/good/satisfactory/not so good/bad' to a single item: 'In general, how would you rate your health?' This type of operationalisation has been shown to be valid and predictive of health indicators in numerous studies (reviewed by Idler and Benyamini¹⁶).

Table 4 shows the results of hierarchical regression analysis for one country, in this case Finland, as an example. To control for potential sociodemographic effects, indicators such as age, sex, education and income were included in the first step of the regression equation. Only age turned out to be a significant predictor of subjective health. From the national indicators used in Finnish PA surveillance so far, two items out of four ('How often did you do leisure-time physical activity with at least light sweating for at least half an hour?' and 'How much exercise or physical exertion in leisure time?') were significant. By including IPAQ-S7T items in the third step,

age and one item of the Finnish national indicators (exercise in leisure time) remained significant predictors. In addition, four out of nine IPAQ indicators were significant. Finally, by including psychosocial and environmental indicators at the fourth step, the one national item on intensity of PA, about exercise or exertion in leisure time (response options ranged from 'little movement' to 'training for sports competition'), turned out to be the strongest behavioural predictor ($\beta = -0.278$) followed by three IPAQ-S7T indicators, the most important being duration of sitting. However, the strongest of all indicators tested was a psychosocial one, i.e. an item about self-efficacy related to PA ('How certain are you that you could do 30 minutes of moderate physical activity if you were sad or tired', responses 5 = 'I'm sure I could' to 1 = 'I'm sure I could not'; $\beta = 0.301$). In addition, one environmental and policy-oriented indicator ('My municipality/city does enough for its citizens concerning their physical activities', responses 5 = 'definitely true' to 1 = 'not true at all') was found to be a significant predictor of subjective health. The particular predictive power of psychosocial and environmental determinants is underlined by the R^2 changes in the different steps of the hierarchical regression analysis. Including the respective indicators at step 4 in the hierarchical regression procedure increased the explained variance of the overall model from about 20% to about 33% (13% increase, significant F change of 4.66).

The results of the regression analysis presented here for Finland have also been observed for other countries participating. For example, for Germany, The Netherlands and the UK, indicators of self-efficacy ('I could do physical activity if sad or tired ... most days a week') turned out to be the strongest predictor of subjective health. Moreover, in most countries R^2 changes were highest from step 3 to step 4, i.e. when including the psychosocial and environmental indicators related to PA.

Discussion

The EUPASS project made major efforts to standardise sampling procedures and fieldwork in the participating countries. Thus, on one hand, the huge differences in response rates may indicate specific challenges for conducting telephone or mail surveys in different EU countries. This should be considered in the further process of developing a European health monitoring system. On the other hand, despite the strenuous efforts of standardisation, remaining differences in procedures should not be underplayed (for a detailed discussion, see Rzewnicki *et al.*⁹). Moreover, as the actual response rates from different countries are rather low, the results of the indicator analyses have to be interpreted with caution. However, for the explorative purposes of the study, the current data appeared to be sufficient.

Table 4 Subjective health status, regressed on different sets of indicators (Finland)

Step/indicators/predictor	Subjective health status (1 = very good, 5 = bad)					
	β t	R ²	F change	β t	R ²	F change
1. Demographic	0.200**	0.052	4.27**	0.152*	0.164*	0.125(*)
				-0.107(*)		
2. National				0.125(*)	0.164	6.87***
				-0.273***	-0.301***	
3. IPAQ-S7T					0.172*	0.199
					-0.127(*)	3.82***
					0.184*	
					-0.140*	
4. Psychosocial and environmental						
						0.197*
						-0.185**
						0.301***
						0.150*
						0.329
						4.66***

IPAQ – International Physical Activity Questionnaire; IPAQ-S7T – short (last 7 days) telephone interview IPAQ; PA – physical activity.

t β – standardised correlation coefficient.

(*), $P < 0.10$; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Reliability of tested indicators

Both the generally rather low test–retest reliability scores for the IPAQ-S7T instrument and the reliability scores reported for most indicators among member states in the EUPASS study may refer to particular methodological issues. First of all, the IPAQ-S7T refers explicitly to the last 7 days as the time period to consider when answering the questions. Thus, differences in the frequency or duration of PA found for one respondent between T1 and T2 (i.e. about two weeks later) could reflect real differences in physical activities conducted by this person (e.g. three times vigorously active in the week before T1; once vigorously active in the week before T2). Second, the original English version of the IPAQ-S7T had to be translated into the languages of the participating countries. This also may have influenced the understanding of single questions in some countries but would not explain the rather low coefficients for the UK, where no translation had to be made. Finally, it should be noted that results of international reliability tests of IPAQ-S7T conducted by the IPAQ group itself showed comparably higher test–retest reliability (see <http://www.ipaq.ki.se>). Since most aspects of the methodology developed and used by the IPAQ group for application of the instrument were also applied in the EUPASS project (e.g. translation procedures as well as statistical procedures for reliability testing), differences in test–retest reliability may also be due to sample issues. For example, in EUPASS the respondents were randomly selected on a nation-wide basis, while in the international tests by the IPAQ group the instruments were given to samples with rather specific geographical and socio-demographic characteristics, including convenience samples.

Comparability of tested indicators

Several EU countries are considering using scientifically tested and internationally comparable indicators such as the IPAQ instrument for their national surveillance efforts. However, neither the European Commission nor the individual EU countries want to lose any information on health data that was collected in the past. Thus, the question of the comparability of old and new indicators is a crucial issue for the conduct of the EUPASS indicator test survey. In sum, the results of the present analysis are not very encouraging regarding the double challenge of necessary change (use more comparable indicators) and desirable continuation (not losing information from data collected in the past). In particular, the comparability of national indicators used to date and the IPAQ-S7T turned out to be very low.

Predictive power of tested indicators

Indicators used in surveillance should be related to major determinants of health. Behavioural determinants such as PA have been considered to be those determinants that are closest to health outcomes. However, as has been

demonstrated by earlier studies^{8,13,15} and in particular by this analysis, psychosocial as well as environmental determinants may have an independent and even larger main effect on subjective health. These results may affect the future design of surveillance systems.

Conclusions

The EUPASS findings underline the need to co-ordinate public health and surveillance activities within the European Community. As has been demonstrated by the comparative investigation of PA indicators in EU countries, the diversity of approaches to measure the population's PA in national surveys is enormous. Existing indicators neither relate to the same concept of health-related PA activity nor focus on comparable dimensions, nor do they apply similar reference periods or scales. As a consequence, available datasets on PA at the country level are not directly comparable at the European level.

One major approach to overcome this situation is related to the efforts of the international consensus group in developing an International Physical Activity Questionnaire. However, before one could recommend EU countries to use the IPAQ as a comparable instrument for national PA surveillance, two key issues have to be dealt with. First of all, the quality and reliability of the IPAQ instrument as well as its international and intercultural applicability and adequacy have to be tested further. For example, as a general result of the EUPASS indicator test survey, reliability coefficients on the short (last 7 days) telephone version (IPAQ-S7T) are generally quite low for all eight participating countries (see Table 3). Second, the comparability of the old indicators and the IPAQ indicators appears to be a particularly important issue for countries such as Finland, which have already had a well-established PA surveillance system since the late 1970s. These countries do not want to lose the possibility of monitoring long-term PA changes in their population over time by substituting their present national indicators with IPAQ indicators, especially if the information provided by the new indicators is no longer comparable to the existing data. In this regard, the results of the current analysis are also not very encouraging. As has been outlined above, the correlation coefficients between old indicators and indicators of IPAQ-S7T are generally quite low in all of the countries investigated in this study.

Countries with a longer tradition in PA monitoring may not wish to substitute their national indicators, but instead to add IPAQ indicators to their system (e.g. Finland, USA). In other cases, such as Belgium (where the main national health survey with two PA items has been conducted only once) or France and Italy (where PA items have been changed and no continuous data are available yet), adoption of an internationally comparable set of indicators (e.g. IPAQ) may be easier.

At least three lessons can be learned from the EUPASS study to guide further activities towards the development of valid, reliable, comparable and health-predictive PA data at the EU level.

1. The value of the data on PA already available at the country level should not be underestimated. For example, as has been shown in Table 4, one of the old indicators used in the national Finnish health survey for many years (PA in leisure time) turned out to be a stronger predictor of subjective health status than the IPAQ-S7T indicators. Of course, this indicator has not been used in other EU countries. Therefore, no internationally comparable data are available. However, new methods of data conversion may be used at the EU level in future to make such national datasets internationally comparable as well¹⁷.
2. The validity and reliability tests of the IPAQ instrument conducted by the IPAQ group itself have provided more promising results than the EUPASS study. As has been mentioned before, only one of the different IPAQ versions has been tested in EUPASS. Further refinement may help to overcome deficiencies shown by the reliability tests in the current study.
3. The importance of psychosocial and environmental determinants for public health has increasingly been recognised in the last few years. In this context, the EUPASS indicator test survey investigated the predictive power of such indicators on subjective health in comparison to old and new (IPAQ) behavioural indicators. As a main result, such psychosocial and environmental indicators turned out to be stronger predictors of subjective health than their behavioural counterparts. Thus, health surveillance that is especially interested in creating a data system for changing public health should at least include a focus on psychosocial and environmental determinants in the future. A European health monitoring system also may consider using such indicators related to self-efficacy and opportunities for PA as tested in the present study.

Acknowledgements

EUPASS was funded by the Health Monitoring Programme of the EU (European Commission, Luxembourg, agreement number VS1999/5133 (99CVF3-502)). The present paper is based on the EUPASS Final Report as delivered to the European Commission and a larger paper on the EUPASS contribution to the European Health Monitoring System, which will be published elsewhere⁶.

Also, the authors wish to acknowledge the support of the following individuals for their contribution to the realisation of this research: I. De Bourdeaudhuij, H. Chamouillet, P. Ciddo, A. Gatrell, P. Godin, A. Haase,

R. Ireland, L. Kannas, M. Lejeune, T. von Lengerke, D. McQueen, J.M. Oppert, N. Rodriguez Avila, J. Rodriguez Diaz, R. Rzewnicki, K. Schmitt, M. Stiggelbout and J. Vinck.

References

- 1 Bouchard C, Shepard R, Stephens T, eds. *Physical Activity, Fitness and Health*. Champaign, IL: Human Kinetics, 1994.
- 2 Blair SN, Booth M, Gyarfás I, Iwane H, Marti B, Matsudo V, *et al.* Development of public policy and physical activity initiatives internationally. *Sports Medicine* 1996; **3**: 157–63.
- 3 US Department of Health and Human Services. *Physical Activity and Health: A Report of the Surgeon General*. Atlanta, GA: Centers for Disease Control and Prevention, 1996.
- 4 Vuori I, Oja P, Stahl T. *Promoting of Health-enhancing Physical Activity. A Preparatory European Meeting*. Tampere, Finland: UKK Institute, 1996.
- 5 King AC. How to promote physical activity in a community: research experiences from the US highlighting different community approaches. *Patient Education and Counselling* 1995; **85**: 1207–11.
- 6 McQueen D, Puska P, eds. *Global Behavioral Risk Factor Surveillance*. New York: Kluwer, in press.
- 7 McQueen D. A world behaving badly: the global challenge for behavioral surveillance. *American Journal of Public Health* 1999; **89**: 1312–4.
- 8 Rütten A, Ziemainz H, Schena F, Stahl T, Stiggelbout M, Vanden Auweele Y, *et al.* Using different physical activity measurements in eight European countries. Results of the European Physical Activity Surveillance System (EUPASS) time series survey. *Public Health Nutrition* 2003; **6**: 371–6.
- 9 Rzewnicki R, De Bourdeaudhuij I, Stahl T, Welshman J, Vanden Auweele Y, Ziemainz H. How methods affect measures of physical activity in eight European countries: the Eupass experience. *Public Health Nutrition* submitted for publication.
- 10 Sallis JF, Owen N. *Physical Activity and Behavioral Medicine*. London: Sage, 1999.
- 11 De Bourdeaudhuij I, Van Oost P, Mommerency G. Daily physical activity in adolescents and young adults. *Archives of Public Health* 1993; **51**(9–10): 407–24.
- 12 De Bourdeaudhuij I, Van Oost P. Differences in level and determinants of leisure-time physical activity between men and women in 3 population-based samples. *Archives of Public Health* 1994; **52**(1–2): 21–45.
- 13 Rütten A, Lüschen G, von Lengerke T, Abel T, Kannas L, Rodriguez Diaz JA. *Health Promotion Policy in Europe. Rationality, Impact, and Evaluation*. Munich: Oldenbourg, 2000.
- 14 Rütten A, Abel T, Kannas L, von Lengerke T, Luschen G, Diaz JA, *et al.* Self reported physical activity, public health, and environment: results from a comparative European study. *Journal of Epidemiology and Community Health* 2001; **55**: 139–46.
- 15 Stahl T, Rütten A, Nutbeam D, Bauman A, Kannas L, Abel T, *et al.* The importance of the social environment for physically active lifestyle – results from an international study. *Social Science and Medicine* 2001; **52**: 1–10.
- 16 Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior* 1997; **38**: 21–37.
- 17 Van Buuren S, Eyres S, Tennant A, Hopman-Rock M. *Response Conversion: A New Technology for Comparing Existing Health Information*. Leiden, The Netherlands: TNO Prevention and Health, 2001.