

Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) - Short Form,

Version 2.0. April 2004

Introduction

This document provides an revision to the outline for scoring the short form of the International Physical Activity Questionnaire (IPAQ) . This is available on the website www.ipaq.ki.se.

There are many different ways to analyse physical activity data, but to-date there is no consensus on a 'correct' method for defining or describing levels of activity based on self-report surveys. The use of different scoring protocols makes it very difficult to compare within and between countries, even when the same instrument has been used.

IPAQ is an instrument designed primarily for population surveillance of adults. It has been developed and tested for use in adults (age range of 15-69 years) and until further development and testing is undertaken the use of IPAQ with older and younger age groups is not recommended. IPAQ is being used also as an evaluation tool in some intervention studies, but the range of domains and types of activities included in IPAQ should be carefully noted before using it in this context.

This document describes the *April 2004 revision* to the IPAQ short scoring protocol¹. These revisions have been suggested by the IPAQ scientific group, to examine variation among countries in more detail². Given the broad range of domains of physical activity asked in IPAQ, new cutpoints need to be trialed and developed to express physical activity in the population. These cutpoints are preliminary, in the sense that they are not yet supported by epidemiological studies, which have typically used Leisure time physical activity (LTPA) to examine benefits or risks of being active. Hence, "30 minutes of moderate intensity PA on most days of the week" was evidence-based, using the estimates of risk (reduction) from these LTPA measures in numerous epidemiological studies.

A new set of suggested cutpoints is based on work in the area of total physical activity, specifically total walking, where recommendations of at least 10,000 steps, and possibly 12,500 steps per day are considered 'high active' (Tudor Locke reference). This equates to at least 2 hours of all forms of walking per day, which includes all settings and domains of activity, and could be a population goal for total HEPA (health-enhancing physical activity). With this background, new cutpoints are proposed for expressing physical activity levels in populations using generic physical activity measures such as IPAQ³.

¹ The first version of an IPAQ scoring protocol was in August 2003; this is a revised version, April 2004. This revised version does not change the continuous forms of reporting data, but does suggest a new category for describing the most active groups in populations. The changes from the August 2003 scoring protocol are indicated in this document.

² Previous scoring algorithms returned high prevalence rates with limited variation among countries; hence a higher cutpoint is sought, as the IPAQ instrument measures total PA, including LTPA as well as incidental, occupational and transport related PA all in one question. This results in much higher prevalence estimates than measures of LTPA alone.

³ This results in changes to the categories used for levels of activity, and to the truncation rules [as greater than two hours per day may be required as usable data for walking and other physical activity behaviors].

Characteristics of the IPAQ short-form instrument:

- 1) IPAQ assesses physical activity undertaken across a comprehensive set of domains including leisure time, domestic and gardening (yard) activities, work-related and transport-related activity;
- 2) The IPAQ short form asks about three specific types of activity undertaken in the three domains introduced above and sitting. The specific types of activity that are assessed are walking, moderate-intensity activities and vigorous intensity activities; frequency (measured in days per week) and duration (time per day) are collected separately for each specific type of activity.
- 3) The items were structured to provide separate scores on walking; moderate-intensity; and vigorous-intensity activity as well as a combined total score to describe overall level of activity. Computation of the total score requires summation of the duration (in minutes) and frequency (days) of walking, moderate-intensity and vigorous-intensity activity;
- 4) Another measure of volume of activity can be computed by weighting each type of activity by its energy requirements defined in METS (METs are multiples of the resting metabolic rate) to yield a score in MET–minutes. A MET-minute is computed by multiplying the MET score by the minutes performed. MET-minute scores are equivalent to kilocalories for a 60 kilogram person. Kilocalories may be computed from MET-minutes using the following equation: MET-min x (weight in kilograms/60 kilograms). The selected MET values were derived from work undertaken during the IPAQ Reliability Study undertaken in 2000-2001. Using the Ainsworth et al. Compendium (*Med Sci Sports Med* 2000) an average MET score was derived for each type of activity. For example; all types of walking were included and an average MET value for walking was created. The same procedure was undertaken for moderate-intensity activities and vigorous-intensity activities. These following values continue to be used for the analysis of IPAQ data: Walking = 3.3 METs, Moderate PA = 4.0 METs and Vigorous PA = 8.0 METs⁴.

Analysis of IPAQ

Both categorical and continuous indicators of physical activity are possible from the IPAQ short form. However, given the non-normal distribution of energy expenditure in many populations, the continuous indicator is presented as median minutes or median MET–minutes rather than mean minutes or mean MET-minutes.

Categorical score

Regular participation is a key concept included in current public health guidelines for physical activity.⁵ Therefore, both the total volume and the number of day/sessions are included in the IPAQ analysis algorithms. There are three levels of physical activity suggested for classifying

⁴ Note that there is still some debate about whether 8 Mets for vigorous is sustainable, in occupational settings for several hours; we have no data on this, but it is likely to be less than that, maybe 7 METs or even less; however, for the moment, we suggest keeping with the compendium value of * METs.

⁵ Pate RR, Pratt M, Blair SN, Haskell WL, Macera CA, Bouchard C et al. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *Journal of American Medical Association* 1995; 273(5):402-7. and U.S. Department of Health and Human Services. *Physical Activity and Health: A Report of the Surgeon General*. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, The Presidents' Council on Physical Fitness and Sports: Atlanta, GA:USA. 1996.

populations; these are the new proposed levels, which take account of the concept of total physical activity of all domains. The proposed levels are:

- [i] 'inactive
- [ii] 'minimally active'⁶
- [iii] 'HEPA active' (health enhancing physical activity; a high active category).

The criteria for these three levels are shown below.

1. Inactive (CATEGORY 1)

This is the lowest level of physical activity. Those individuals who not meet criteria for Categories 2 or 3 are considered '**insufficiently active**' [CATEGORY 1].

2. Minimally Active (CATEGORY 2)

The minimum pattern of activity to be classified as 'sufficiently active' is any one of the following 3 criteria:

- a) 3 or more days of vigorous activity of at least 20 minutes per day **OR**
- b) 5 or more days of moderate-intensity activity or walking of at least 30 minutes per day **OR**
- c) 5 or more days of any combination of walking, moderate-intensity or vigorous intensity activities achieving a minimum of at least 600 MET-min/week.

Individuals meeting at least one of the above criteria would be defined as achieving the minimum recommended to be considered '**minimally active**' [CATEGORY 2]. This category is more than the minimum level of activity recommended for adults in current public health recommendations, but is not enough for "total PA" when all domains are considered. IPAQ measures total physical activity whereas the recommendations are based on activity (usually leisure-time or recreational) over and above usual daily activities.

3. HEPA active (CATEGORY 3)

A separate category labeled '**HEPA**' level, which is a more active category [CATEGORY 3] can be computed for people who exceed the minimum public health physical activity recommendations, and are accumulating enough activity for a healthy lifestyle. This is a useful indicator because it is known that higher levels of participation can provide greater health benefits, although there is no consensus on the exact amount of activity for maximal benefit. Also, in considering lifestyle physical activity, this is a total volume of being active which reflects a healthy lifestyle. It is at least 1.5 – 2 hours of 'being active' throughout the day, which is more than the LTPA-based recommendations of 30 minutes⁷.

In the absence of any established criteria, the IPAQ scientific group proposes this new cutpoint, which equates to approximately at least 1.5 -2 hours of total activity per day, of at least moderate-intensity activity. It is desirable to have a 'HEPA' activity category, because in some populations, a large proportion of the population may be classified as "minimally active" because the IPAQ instrument assess all domains of activity. Category 3 sets a higher threshold of activity and provides a useful mechanism to distinguish variation in sub-population groups.

⁶ "Minimally active" implies some physical activity but is not an optimal level of total HEPA.

⁷ As Tudor-Locke and others have indicated, there is a basal level of around 1 hour of activity just in activity of daily living, and an additional 0.5 – 1 hour of LTPA makes a healthy lifestyle amount of total PA – hence, these new cutpoints are still consistent with the general LTPA based public health recommendations of at least half an hour per day of additional activity or exercise.

The two criteria for classification as ‘HEPA active’ are:

- a) vigorous-intensity activity on at least 3 days achieving a minimum of at least 1500 MET-minutes/week **OR**
- b) 7 or more days of any combination of walking, moderate-intensity or vigorous intensity activities achieving a minimum of at least 3000 MET-minutes/week⁸

Continuous score

Data collected with IPAQ can be reported as a continuous measure and reported as median MET-minutes. Median values can be computed for walking (W), moderate-intensity activities (M), and vigorous-intensity activities (V) using the following formulas:

MET values and Formula for computation of Met-minutes

Walking MET-minutes/week = 3.3 * walking minutes * walking ‘days’

Moderate MET-minutes/week = 4.0 * moderate-intensity activity minutes * moderate days

Vigorous MET-minutes/week = 8.0 * vigorous-intensity activity minutes * vigorous-intensity days

A combined total physical activity MET-min/week can be computed as the sum of Walking + Moderate + Vigorous MET-min/week scores.

The MET values used in the above formula were derived from the IPAQ validity and reliability study undertaken in 2000-2001⁹. A brief summary of the method is provided above (see page 1).

As there are no established thresholds for presenting MET-minutes, the IPAQ Research Committee proposes that these data are reported as comparisons of median values and interquartile ranges for different populations.

IPAQ Sitting Question

The IPAQ sitting question is an additional indicator variable and is not included as part of any summary score of physical activity. Data on sitting should be reported as median values and interquartile range. To-date there are few data on sedentary (sitting) behaviors and no well-accepted thresholds for data presented as categorical levels.

Data Processing Rules

In addition to a standardized approach to computing categorical and continuous measures of physical activity, it is necessary to undertake standard methods for the cleaning and treatment of IPAQ datasets. The use of different approaches and rules would introduce variability and reduce the comparability of data.

There are no established rules for data cleaning and processing on physical activity. Thus, to allow more accurate comparisons across studies IPAQ has established and recommends the following guidelines:

1. Data cleaning

- time should be converted from hours and minutes into minutes

⁸ Note: this replaces the previous IPAQ short form cutpoint of 1500 mets.mins/ week

⁹ Craig CL, Marshall A, Sjostrom M et al. International Physical Activity Questionnaire: 12 country reliability and validity Med Sci Sports Exerc 2003;August.

- ensure that responses in ‘minutes’ were not entered in the ‘hours’ column by mistake during self-completion or during data entry process, values of ‘15’, ‘30’, ‘45’, ‘60’ and ‘90’ in the ‘hours’ column should be converted to ‘15’, ‘30’, ‘45’, ‘60’ and ‘90’ minutes, respectively, in the minutes column.
- time should be converted to daily time [usually is reported as daily time, but a few cases will be reported as optional weekly time – eg. VWHRS, VWMINS – convert to daily time]
- convert time to mets-mins [see above; days x daily time]
- must have the number of days for the day variables; for the ‘time’ variables, either daily or weekly time is needed – if ‘don’t know’ or ‘refused’ or data are missing in walking, moderate or vigorous days or minutes, then that case is removed from analysis

2. Maximum Values for excluding outliers

This rule is to exclude data which are unreasonably high; these data are to be considered outliers and thus are excluded from analysis. All Walking, Moderate and Vigorous time variables which total at least or greater than ‘16 hours’ should be excluded from the analysis.

The ‘days’ variables can take the range 0-7 days, or 8,9 (don’t know or refused); values greater than 9 should not be allowed and those data excluded from analysis.

3. Truncation of data rules

This rule is concerned with data truncation and attempts to normalize the distribution of levels of activity which are usually skewed in national or large population data sets. It is recommended that all Walking, Moderate and Vigorous time variables exceeding ‘4 hours’ or ‘240 minutes’ are truncated (that is re-coded) to be equal to ‘240 minutes’ in a new variable¹⁰. This rule permits a maximum of 28 hours of activity in a week to be reported for each category of physical activity. *This rule requires further testing, but is an initial manner proposed for classifying these population data.*

When analysing IPAQ data and presenting the results in categorical variables, this rule has the important effect of preventing misclassification in the ‘high active’ category. For example, an individual who reports walking for 2.5 hours every day and nothing else would be classified as ‘HEPA active’ (reaching the threshold of 7 days, and ≥ 3000 MET.mins. Similarly, someone who reported walking for 90 minutes on 5 days, and 4 hours (240 mins) of moderate activity on another day and 70 minutes of vigorous activity on another day, would also be coded as ‘HEPA active’ because this pattern meets the ‘7 day’ and “3000 MET-min” criteria for ‘HEPA active’.

4. Minimum Values for Duration of Activity

Only values of 10 or more minutes of activity will be included in the calculation of summary scores. The rationale being that the scientific evidence indicates that episodes or bouts of at least 10 minutes are required to achieve health benefits. Responses of less than 10 minutes [and their associated days] should be re-coded to ‘zero’.

Summary of Data Processing Rules 1- 4 above

Data management rules 2, 3, and 4 deal with first excluding outlier data, then secondly, recoding high values to ‘4 hours’, and finally describing minimum amounts of activity to be included in

¹⁰ Note that this is a different truncation rule to the earlier scoring protocol; we have previously used 2 hours as a truncation point for LTPA measures. This higher truncation point is proposed in order to allow people who walk for 2.5 hours per day and do nothing else to be categorized as ‘HEPA’ active; if data were truncated, these individuals would be recoded to 2 hours per day, and over 7 days, total 2772 MET.mins, due to the truncation rule. The new truncation rule allows 2.5 hours to be counted in full. The initial purpose of truncation was to normalize the distributions, and was based on expert judgments. It is now suggested that 4 hours / day be proposed as a truncation threshold for more inclusive ‘lifestyle PA measures’ such as IPAQ.

analyses. These rules will ensure that highly active people remain highly active, while decreasing the chances that less active individuals are coded as highly active.

5. Calculating Total Days for ‘minimally Active’ [category 2] and ‘HEPA Active’ [category 3]

Presenting IPAQ data using categorical variables requires the total number of ‘days’ on which all physical activity was undertaken to be assessed. This is difficult because frequency in ‘days’ is asked separately for walking, moderate-intensity and vigorous-intensity activity, thus allowing the total number of ‘days’ to range from a minimum of 0 to a maximum of 21 ‘days’ per week. The IPAQ instrument does not record if different types of activity are undertaken on the same day.

In calculating ‘**minimal activity**’, the primary requirement is to identify those individuals who undertake a combination of walking and/or moderate-intensity activity on at least ‘5 days’/week. Individuals who meet this criterion should be coded in a new variable called “*at least five days*”.

Below are two examples showing this coding in practice:

- i) an individual who reports ‘2 days of moderate’ and ‘3 days of walking’ should be coded as a value indicating “*at least five days*”;
- ii) an individual reporting ‘2 days of vigorous’, ‘2 days walking’ and ‘2 days moderate’ should be coded as a value to indicate “*at least five days*” [even though the actual total is 6].

The original frequency of ‘days’ for each type of activity should remain in the data file for use in the other calculations.

The same approach as described above is used to calculate total days for computing the ‘**HEPA active**’ category. The primary requirement according to the stated criteria is to identify those individuals who undertake a combination of walking, moderate-intensity and or vigorous activity on at least 7 days/week. Individuals who meet this criterion should be coded in a value in a new variable to reflect “*at least 7 days*”.

Below are two examples showing this coding in practice:

- i) an individual who reports ‘4 days of moderate’ and ‘3 days of walking’ should be coded as the new variable “*at least 7 days*”.
- ii) an individual reporting ‘3 days of vigorous’, ‘3 days walking’ and ‘3 days moderate’ should be coded as “*at least 7 days*” [even though the total adds to 9].

Summary: The algorithm(s) in Appendix 1 and Appendix 2 to this document show how these rules work in an analysis plan, to develop the categories 1 [inactive], 2 [minimally], and 3 [HEPA] levels of activity. A short form [‘at a glance’] and a diagram showing these analytic steps for ‘sufficient physical activity’ and ‘high active’ categories are shown as appendix 1 at the end of this document.

*IPAQ Research Committee
April 2004*

APPENDIX 1

**At A Glance
IPAQ Scoring Protocol (Short Versions)**

Categorical Score- three levels of physical activity are proposed

1. Inactive

- No activity is reported **OR**
- Some activity is reported but not enough to meet Categories 2 or 3.

2. Minimally Active

Any one of the following 3 criteria

- 3 or more days of vigorous activity of at least 20 minutes per day **OR**
- 5 or more days of moderate-intensity activity or walking of at least 30 minutes per day **OR**
- 5 or more days of any combination of walking, moderate-intensity or vigorous intensity activities achieving a minimum of at least 600 MET-min/week.

3. HEPA active

Any one of the following 2 criteria

- Vigorous-intensity activity on at least 3 days and accumulating at least 1500 MET-minutes/week **OR**
- 7 or more days of any combination of walking, moderate-intensity or vigorous intensity activities achieving a minimum of at least 3000 MET-minutes/week

Continuous Score

Expressed as MET-min per week: MET level x minutes of activity x events per week

Sample Calculation

MET levels	MET-min/week for 30 min episodes, 5 times/week
Walking = 3.3 METs	3.3*30*5 = 495 MET-min/week
Moderate Intensity = 4.0 METs	4.0*30*5 = 600 MET-min/week
Vigorous Intensity = 8.0 METs	8.0*30*5 = 1,200 MET-min/week
	TOTAL = 2,295 MET-min/week

Total MET-min/week = (Walk METs*min*days) + (Mod METs*min*days) + Vig METs*min*days)

Please review the document “Guidelines for the data processing and analysis of the International Physical Activity Questionnaire (Short Form)” for more detailed description of IPAQ analysis and recommendations for data cleaning and processing [www.ipaq.ki.se].

START HERE

