Criterion validity and test-retest reliability of a physical activity questionnaire in South African primary school-aged children

Joanne A McVeigh, Shane A Norris

Exercise Laboratory, School of Physiology, and MRC/WITS Developmental Pathways for Health Research Unit, Department of Paediatrics, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

PhD Shane A Norris, PhD

MRC/WITS Developmental Pathways for Health Research Unit, Department of Paediatrics, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Corresponding author: J A McVeigh (jo-anne.mcveigh@wits.ac.za).

Abstract

Objective. We sought to determine the validity, reliability and ranking ability of an interviewer-administered physical activity questionnaire (PAQ), measuring physical activity (PA) and inactivity history over a 1-week and 1-year period in South African primary school-aged children.

Methods. Criterion validity of the PAQ was tested against PA movement counts as measured with an Actical accelerometer in 30 children. Agreement between the two instruments was measured with a weighted Kappa statistic. Test-retest reliability of the past week and past year PAQ was also tested.

Results. A positive, significant ($r=0.53, p=0.004$) relationship was found between total time spent being physically active as measured by the Actical and PAQ. A similar relationship was found for time spent doing sedentary ($r=0.63, p<0.001$) and vigorous activities ($r=0.47, p<0.001$), but not for activities of a moderate intensity ($r=0.01, p=0.88$). The ability of the PAQ to correctly categorise children into activity levels was moderate ($k=0.41, p<0.001$). The PAQ was found to be reliable and reproducible with significant ($p<0.001$) intraclass correlation coefficients for both the past week and past year administrations.

Conclusion. The interviewer-administered PAQ is a useful assessment tool in this population of children, as evidenced by its good correlation with Actical measurements. The best application of the questionnaire lies in its ability to qualitatively rank subjects according to activity level.


Introduction

South African children are showing trends of obesity and overweight, similar to values reported from high-income countries 15 years ago. The decline in overall physical activity (PA) levels has occurred concomitantly with the rising trend of obesity, and less than one-third of South African children now participate in sufficient (defined as participation in activities such as soccer, netball, rugby, basketball or running for 20 minutes or more on at least 3 of the 7 days preceding the survey) PA on a weekly basis. There is an ethnic disparity regarding levels of PA in South African children, with a greater percentage of black youths (37.5%) found to be insufficiently active compared with white youths (29.4%). Moreover, less than one-third of 9-year-old black South African children report having structured physical education classes at school. This coupled with evidence for declining PA levels with increasing age is concerning.

Physical activity intervention strategies aimed at improving an increasingly sedentary lifestyle need to be developed for South African children. To do this, we need accurate, reliable measurements of PA, which are feasible for use within a South African context. Characteristics of physical activity questionnaires (PAQs) that favour their use in studies with large sample sizes include the low cost of producing and administering the questionnaires, the simplicity of administering the questionnaire, their applicability of use in any setting and their validity in terms of measuring activity. On the other hand, PAQs also have limitations. Accurate assessment is reliant on a person’s ability to correctly recall and record their PA undertaken in the preceding period. In addition, there is large variability in the design of PAQs and many have been found to have varying rates of validity.

International validation studies performed in children and adolescents that have compared PAQs with heart rate monitoring, accelerometry and other PA measures have in general shown good validity of PAQs. However, there is a lack of locally validated questionnaires for the purpose of measuring and reporting PA levels in South African children. While there have been studies conducted within African populations which have used international physical activity recall questionnaires (previous day physical activity recall questionnaire (PDPAR)) and heart rate monitoring as an objective comparative measure, few studies have actually objectively validated a PAQ specifically in South African children.
In the study of Mciza and colleagues, the authors made use of an ACTIVITYGRAM as their comparative measure against a PAQ. The authors reported weak but significantly positive associations between the two measures. Prista and colleagues have recently used accelerometry to objectively measure physical activity levels in children living in rural Mozambique. To our knowledge, no other study has used actigraphy to validate a PAQ in South African children. The PAQ used in the current study is based on questionnaires which have been previously validated. While construct validity of the PAQ has been established in large cohorts of South African children, the main aim of our study was to determine criterion validity of the PAQ. We examined whether a subjective method of estimating PA and inactivity levels in South African children (using an interviewer-administered PAQ is valid, reliable and useful for classifying children into categories of physical activity (high to low) by comparing it with an objective measure of PA (Actical accelerometer).

Methods

Subjects

Data were collected from a convenience sample of children recruited from an English-speaking public primary school in Johannesburg, South Africa. Thirty black and white children aged between 9 and 11 years of age participated in a 4-day Actical and PAQ-based study. Based on previously published correlation coefficients between PAQs and Actical data which range between 0.16 and 0.58, we chose a value of 0.53 (moderate strength) and conducted a sample size calculation on Statistica v10. It was estimated that a sample size of 30 children was needed at a power of 90% to ascertain a correlation coefficient of this size. Children who were unable to participate in PA were excluded from the study. All children and caregivers gave written informed consent. The study was approved by the Human Ethics Committee of the University of the Witwatersrand (ethics no: MO50226).

Anthropometric measures

The height of each child, recorded to the nearest millimetre, was measured using a stadiometer (Holtain, UK), and weight, recorded to the nearest 100 gram, was measured using a digital scale (Dismed, USA). Participants were measured with light clothing and no shoes.

Physical activity questionnaire

The PAQ used in this study is based on questionnaires which have been previously validated. However, some questions were modified to make them more appropriate to assess the activities of South African children. Our questionnaire assessed PA and inactivity over a period of 7 days and over the past year. The questionnaire was interviewer-administered and children were asked questions regarding estimates of time spent doing physical activities over a 1-week and 1-year period.

From the questionnaire, we measured the total time spent in all PA in relation to school, sports (at school and outside of school), commuting to and from school and other leisure time activities (including informal activities, sedentary activities such as television watching, etc.) during the week and on the weekend. A total PA score was calculated from the questionnaire: for all subjects, all activities were rated according to their metabolic intensity based on the method of Ainsworth et al. The metabolic PA score (METPA) was calculated by multiplying the intensity (multiples of basal metabolic rate (metabolic equivalents)) by the duration of the activity (hrs/wk). A sum score METPA was calculated as the sum of all METPA scores for each activity. This METPA score is a gross estimate of activity since we did not subtract the resting energy expenditure (one MET) from the gross cost of each activity, as has been used previously by Haapanen et al.

Objective measure of physical activity: accelerometry

In accordance with international best practice guidelines, participants wore Actical (Mini Mitter Co., Inc., Bend, OR) accelerometers (hereafter referred to as Acticals) affixed on their left hip, slightly above the iliac crest, with an elastic belt for 4 consecutive days. Acticals were positioned in this way so that acceleration of the displacement of the hip could be measured and so that the Actical would also be sensitive to weight-bearing movements. When positioned on the hip, the device is extremely sensitive to vertical movements of the torso. Activity was measured in 1-minute epochs (the recommended epoch length) over a period of 4 days. In this study, 60 1-second values were summed together to generate one resultant raw activity datum (counts) for each 1-minute epoch. The child’s age, weight, height and gender were entered into the Acticals memory so that total energy expenditure and PA counts could be estimated using the manufacturer’s equation. The Actical is activated when the participant’s information is entered into the unit via a computer interface, and inactivated when the information is downloaded to a computer. A new battery was fitted to each Actical before activation.

Children’s physical activities were divided into three categories based on the number of activity counts generated by the Actical: sedentary, moderate and vigorous activity. The cut points described by Puyau et al. were used. Sedentary and light activities had activity counts of less than 1 499, moderate activities had counts of between 1 500 and 6 499 and counts greater than 6 500 were classified as vigorous activities.

Study design

Data collection took place over a 2-week period, as depicted in Fig. 1. The past week PAQ was administered on the Thursday prior to the Actical assessment. This questionnaire examined all activity and inactivity that subjects had undertaken in the past 7 days. The past year PAQ was administered the next day (Friday – day 1 of the Actical assessment) and examined all activity and inactivity that subjects had undertaken in the past year. The Actical assessment period then took place over 4 days, from and including a Friday, Saturday, Sunday and Monday, such that data from both weekday and weekend activity were obtained. The data collected from the Acticals were then downloaded onto a computer program (Actical v2.000.7 by Mini Mitter Co., Inc) and comparisons were made with data obtained from the PAQs. Children were asked to behave as they normally would for the 4 days of the study and to wear the Actical at all times except when bathing or showering. All subjects wore their Acticals continuously (except during showering or bathing), as evidenced by the continuous data and counts obtained when the data were downloaded.

The accelerometers were brought to the child’s classroom and collected again 4 days later. All procedures were explained to study participants in detail; they were shown how to attach the Actical to their hips and the PAQ was interviewer administered on all occasions. The past week PAQ was administered for a second time 1 week after the first administration (following Thursday) in order
to assess whether the week during which subjects wore the Actical was representative of a normal active week for the subjects. The past year PAQ was administered again the next day (Friday) to ascertain whether the data obtained on the past year questionnaire were reproducible.

Data analysis
All data were analysed using SPSS for windows v12.0 (SPSS Inc, Chicago, IL, USA). Data are presented as means (SD) unless otherwise indicated. Intra-class correlation coefficients (ICCs) (one-way random effects model) were utilised to evaluate the reproducibility of the estimates of PA obtained from the PAQ questionnaire. Pearson product-moment correlations assessed the magnitude of the relationship between Actical measurements and those reported by the subjects on the PAQ. In addition, the level of agreement between the PAQ and Actical measurement for time spent in sedentary and vigorous activity was estimated using Bland-Altman plots. Weighted Kappa statistics were used to assess how accurate the PAQ is at categorising children into various activity levels. The weighted Kappa statistic provides a measure of agreement between two scores which classify observations into one of several groups or categories. Kappa values were defined as poor (<0.20), fair (0.21 - 0.40), moderate (0.41 - 0.60) and good (0.61 - 1.00). An alpha of 0.05 was set a priori to determine the level of statistical significance.

Results
Sample characteristics
Thirty children participated in this study and had a mean ± SD age of 10.17±0.80 years. Approximately half of the subjects were black (47%) and 60% of the subjects were female. The mean height of the subjects was 139.2±9.0 cm and weight was 35.2±8.4 kg. Average activity counts and total METPA score obtained over the 4-day study period were 256±55 Cal/day over the 4-day study period. Although children had slightly lower average energy expenditure (1 204±54 Cal/day) over the weekday period (Friday and Monday) compared with the weekend period (1 306±57 Cal/day Saturday and Sunday), there was no significant difference between these two time frames in this sample of children.

Validity
An overall significant, positive correlation was observed for total activity counts and total METPA score obtained over the 4-day period. The correlation coefficients are shown in Table 1.
period ($r=0.53, p=0.004$) for all children. To determine whether the relationship held true for activities of varying intensities, correlation coefficients were calculated for time spent in sedentary, moderate and vigorous activity as reported by the subjects on the PAQ (for the same 4-day period as the Actical) and as measured by the Actical. A large, positive ($r=0.63$) and significant ($p=0.0003$) correlation was observed between the average percentage time spent per day in sedentary activities as reported on the PAQ and the average percentage of time per day spent in sedentary activity as measured by the Actical. Both measures included time spent asleep.

Although the means of the PAQ and Actical were similar for activities of moderate intensity, there was no correlation between the two measures for moderate intensity ($r=0.001, p=0.88$). For activities of a high intensity, a medium, positive ($r=0.47$) and significant ($p=0.011$) relationship between the average time spent per day in vigorous activity (METPA score $>6$) and the percentage of time in activities classified as vigorous by the Actical (activity counts $>6500$) was observed.

To assess agreement between the two measures, Bland-Altman comparisons on the time spent in sedentary, moderate and vigorous activity were done. Figs 2, 3 and 4 show bias and variability between the two methods for low, medium and high intensity activities. The PAQ demonstrated good agreement with the Actical for activities of low intensity. Activities of moderate intensity were not well agreed upon by the two measures although similar means were found. Despite the moderate correlation observed for activities of a higher intensity, the PAQ tends to overestimate time spent in vigorous activity (Fig. 4).

A weighted Kappa statistic (which assumes the categories to be ordered) assessed the strength of the PAQ at placing children into low, moderate or high physical activity groups. The PAQ has a moderate ability ($κ=0.41$) to correctly categorise children into physical activity groups.

**Discussion**

The primary purpose of this exploratory study was to make use of Actical accelerometry to determine the validity and reliability of a PAQ in healthy South African children. We wanted to determine whether an interviewer-administered PAQ could be used to accurately assess whether a child can be classified as highly, moderately or not very physically active in relation to their peers. The PAQ demonstrated acceptable levels of reliability and validity when compared with objective data obtained from the Actical accelerometer for activities of a low and high intensity, but not for moderate intensity activities. Despite the moderate to strong validity coefficients for sedentary and vigorous activities, the objective data obtained from the Actical indicate that the PAQ is better at approximating time spent in sedentary activity and overestimates time spent in vigorous PA. The strength of the PAQ used in this study is its ability to qualitatively rank subjects from high to low overall categories of PA. The PAQ has a moderate ability to correctly rank children into varying levels of objectively determined physical activity levels by the Actical.

Although this study provides valuable information which should be taken into account when wanting to quantify levels of physical activity in children, there were limitations to the study. Care needs to be taken when using the PAQ for further research. This study made use of a fairly homogenous small sample of children from one school. This may limit the applicability of the results to the wider population of South African school children. Future studies should examine...
the use of Acticals in a wider age and geographical range of South African children. Acticals have limited use in measuring activity that is horizontal in nature. Activities such as cycling may not have been accurately recorded by the Actical. Combining the use of Acticals with heart rate monitoring will help to overcome this limitation in future studies. A further limitation concerns the assignment of adult MET scores to children’s physical activities. However, since there is a paucity of data on the energy costs of children’s activities, Ridley and Olds conducted an extensive review and concluded that there is sufficient evidence that adult MET scores be used to allot energy costs to children as they are currently the best technique available.38 The same authors went on to develop their own compendium of energy expenditure estimates for youth despite reporting that the error associated with using adult MET scores for children is very small.39 We have not used this compendium in this study as we had previously used the PAQ in other studies and wanted to keep the gross MET score allocation consistent. The gold standard for measuring energy expenditure in free living humans is the doubly-labelled water technique, which is expensive and has not been used in humans in South Africa. Additionally many questionnaires are specifically designed for use in developed countries and their transfer to other cultures in developing countries may not always be appropriate.40 Using accelerometry to objectively monitor physical activity is a more costly alternative to questionnaires. Acticals are an expensive item to purchase, particularly in South Africa, and may therefore be suitable for activity monitoring in a small group of children only. The feasibility of using actigraphy in a large cohort of South African children would largely be fund dependent. There are also other less expensive and validated objective tools, such as the Kenz lifecorder and pedometers which have been used in studies outside of South Africa.37-40

To date only one study has designed and validated a questionnaire specific to a population in Africa,41 although the Global Physical Activity Questionnaire (GPAQ) has been used and validated within an adult South African population and was found to be a reliable questionnaire to use to estimate time spent in occupational, transport and leisure time respectively \(r = 0.74; r = 0.74; r = 0.71\).42-46 The International Physical Activity Questionnaire (IPAQ) has also been shown to be a valid and reliable measure of physical activity in an adult South African population,47 but has not yet been used in children. One other study has reported weak but significant correlations between a PAQ and ACTIVITYGRAM in South African girls.48

It was necessary to identify a suitable PAQ to be used in a South African context that has the ability to differentiate very active children from inactive children. A wide range of reliability \((R = 0.51 - 0.98)\) and validity \((r = 0.20 - 0.88)\) coefficients between PAQs and accelerometry has been reported in the literature.44-46 and although these reported ranges are rather wide, the correlations reported in the present study fall well within them and towards the upper end of both ranges. High reproducibility coefficients established that the week in which the PAQ was administered was representative of an average week for the children. It is also possible that the children gave average answers irrespective of what their actual activity was for that week. Acticals are able to discriminate between sedentary, light, moderate and vigorous activity levels.49 In the present study, Actical output was considered as an overall measure of activity and as percentages of time spent within ranges of activity levels (sedentary to vigorous).

Three findings are of importance. Firstly, that the PAQ and the Actical are correlated at a low range of activity; secondly, the relationship between the two instruments holds true for activities performed at a high intensity and thirdly that PA as reported on the PAQ was consistently reported on the weekly and yearly recall of activity by the subjects. Importantly however, the PAQ is not able to accurately discriminate activities of a moderate intensity, possibly due to the nature of the questions asked. The types of activities associated with a moderate intensity MET score may be rather variable in a child’s recall of their activity. Nonetheless when looking at a child’s overall placement in terms of activity in categorical terms and in relation to their peers, the PAQ provided valuable data. It is not surprising that although the two instruments were significantly correlated, the correlation was only moderately strong. It is important to note that our PAQ could result in misclassification of physical activity levels if too narrow a category is used. However, using three or four broader categories ranging from low- to high-activity levels would help prevent incorrect categorisation of children.

The PAQ used in our study is able to differentiate between different intensities of physical activities and is a useful tool for evaluating PA in English-speaking South African children. However, the relationship between the PAQ and Actical assessments of activity is not along the line of identity suggesting that the PAQ overestimates at high values of PA and possibly underestimates at low values. Nevertheless the relationship allows the PAQ to be used to confidently divide children into categories of activity (high to low).

To our knowledge, this is the first study to make use of Actical accelerometry in South African children. Results from the First South African National Youth Risk Behaviour Survey showed that less than one-third of South African children met the recommended requirements of physical activity for children.49 Subsequent findings presented in the Second South African National Youth Risk Behaviour Survey show that this number has indeed worsened, with 41.5% of children surveyed attaining insufficient physical activity levels.5 Our findings are in agreement with this, with only 27% of our participants engaging in the recommended amount of moderate to vigorous activity on the days measured. It is thus imperative that we begin to take seriously the culture of inactivity that is fast becoming the norm among our South African children.

The construct validity of the PAQ already established\(^2,3,24\) is bolstered by the criterion validity findings presented here. The interviewer-administered PAQ was found to be a valid and reliable assessment tool in this population of children, as evidenced by its overall good correlation with Actical measurements. The PAQ is an easy-to-administer, cost-effective tool that can be used for discriminating more active children from those who are less active. The PAQ can be used with ease in large samples of children. While absolute METPA scores can be calculated from the questionnaire, the better application of the questionnaire appears to lie in its ability to qualitatively rank subjects according to activity level.

Acknowledgement. This work was funded by the Medical Research Council (South Africa) and the University of the Witwatersrand.

REFERENCES


