



Validation of the New Zealand Physical Activity Questionnaire (NZPAQ-LF) and the International Physical Activity Questionnaire (IPAQ-LF) with accelerometry

R M Boon, M J Hamlin, G D Steel, J J Ross

Environment Society and Design Division, Lincoln University, Canterbury, New Zealand

Correspondence to

Dr M J Hamlin, Department of Social Science, Parks, Recreation, Tourism and Sport, Lincoln University, PO Box 84, Christchurch, New Zealand; mike.hamlin@lincoln.ac.nz

Accepted 2 October 2008
Published Online First
3 November 2008

ABSTRACT

Background Validation of instruments used to measure physical activity patterns is essential when attempting to assess the effectiveness of physical activity interventions.

Objectives To assess the validity of two self-report physical activity questionnaires on a representative sample of New Zealand adults.

Methods 70 adults aged 18–65 years from around Christchurch, New Zealand were required to wear an ActiGraph GT1M accelerometer during all waking hours for 7 consecutive days. Immediately following the 7 day accelerometer period participants were required to complete the long forms of both the New Zealand Physical Activity Questionnaire (NZPAQ-LF) and the International Physical Activity Questionnaire (IPAQ-LF).

Results Both the NZPAQ-LF and the IPAQ-LF questionnaires showed small to moderate correlations with ActiGraph data for time spent in moderate-intensity physical activity ($r = 0.19$ – 0.30) and total physical activity (sum of moderate and vigorous-intensity physical activity, $r = 0.30$ – 0.32). In comparison with the ActiGraph data, both self-report questionnaires tended to overestimate activity levels by approximately 165%. Total physical activity levels gathered from both questionnaires were strongly correlated with each other ($r = 0.79$) and showed good levels of agreement in the Bland–Altman plots.

Conclusions The long forms of the NZPAQ and IPAQ were found to have acceptable validity when detecting participants' ability to meet activity guidelines based on exercise duration, but a significant amount of over-estimation was evident. This presents a need for both instruments to be further developed and tested in order to increase validity.

It is well known that regular physical activity not only has a positive effect on individuals' fitness levels¹ but is associated with a range of health benefits.² However, recent data suggest that physical activity is not a priority for most people in developed countries, with approximately 74% of American,³ 65% of Canadian,⁴ 69% of British,⁵ 45% of Australian⁶ and 32% of New Zealand adults⁷ not sufficiently active. In order to reverse physical activity and lifestyle disease trends, many countries worldwide have implemented large-scale population-wide physical activity interventions. However, in order to monitor population health and assess the effectiveness of such interventions, accurate measurement of physical activity is necessary. Methods used to quantify habitual

physical activity include objective instruments which measure body movement (e.g. pedometers and accelerometers),^{8–10} or physiological processes (e.g. heart rate monitoring),¹¹ and subjective recall questionnaires.¹² While all instruments have their limitations, the most commonly used method for measurement of population physical activity is the self-report questionnaire.¹³ Self-report questionnaires rely on individuals' understanding and knowledge of the questions posed and their ability to accurately recall and record all physical activity; because of this, such questionnaires often prove to have varying rates of validity, which ultimately affects their usefulness in accurately assessing physical activity levels of populations.

The International Physical Activity Questionnaire (IPAQ) has been developed to estimate levels of habitual physical activity across different countries and socio-cultural environments.¹² The short version of the IPAQ (IPAQ-SF) was developed for surveillance studies and generates less information than the long version (IPAQ-LF), which aims to provide comprehensive information on the duration of moderate and vigorous-intensity physical activity in work, domestic, transportation and leisure-related areas. Sport and Recreation New Zealand (SPARC) and the New Zealand Ministry of Health recently developed two self-report physical activity questionnaires. The New Zealand Physical Activity Questionnaires were designed as either a short version, which was based on the IPAQ-SF and developed as a surveillance tool (NZPAQ-SF), or a long version (NZPAQ-LF), which, unlike the IPAQ-LF, uses a retrospective 7 day diary format to gather detailed physical activity information. Questionnaires can be downloaded from <http://www.ipaq.ki.se> or <http://www.sparc.org.nz>.

Previous research has indicated high reliability for the IPAQ-LF, with an average test–retest correlation coefficient of approximately 0.8 on data from participants in 12 different countries.¹² The test–retest reliability of the NZPAQ-SF was reported to be approximately 0.7; however, reliability data on NZPAQ-LF have not been reported. Although a number of validity studies on the IPAQ questionnaires have been completed, and have reported correlation coefficients of approximately 0.3 with accelerometry¹² and doubly labelled water (DLW),¹⁴ such studies on the validity of the NZPAQ-LF are scarce. Maddison *et al* (2007) reported a correlation coefficient of approximately

0.4 between the NZPAQ-SF and DLW, while the only published validity study on NZPAQ-LF reported moderate correlations (approximately 0.4) between brisk walking and vigorous-intensity physical activity with heart rate monitoring, but a small correlation (<0.1) between moderate-intensity physical activity and heart rate monitoring.¹⁵

The aim of this study was to compare objectively measured physical activity levels gathered using accelerometry with two self-report questionnaires. This study will report physical activity levels of a cross-section of adult New Zealanders, in addition to providing information on the validity of two potentially useful questionnaires that are likely to be utilised in future physical activity studies.

MATERIALS AND METHODS

Seventy participants aged 18–65 years from around the Christchurch region in New Zealand participated in this study, which was conducted from November 2007 to February 2008. Participants were recruited randomly from four different shopping malls located in areas of differing socio-economic status throughout the Christchurch metropolitan area. In order to gain a representative demographic sample, a quota sampling technique was used which was based on the distribution of age bands and sexes in the general New Zealand population. Informed voluntary consent was attained from all participants prior to their inclusion in the research. This study had the approval of the Lincoln University Human Ethics Committee (reference 2007–60).

The ActiGraph GT1M (Shalimar, Florida, USA) is a small ($3.8 \times 3.7 \times 1.8$ cm, 27 g) uniaxial accelerometer that measures acceleration in the vertical direction. The monitor is designed to detect accelerations that occur from normal human motion and disregard high-frequency vibrations that might occur from mechanical equipment. The ActiGraph contains a microprocessor that filters accumulated signals at a rate of 30 Hz and converts the signal to a numeric value known as activity counts. The ActiGraph accelerometer, formerly known as the CSA monitor, is widely used in physical activity research and has been shown to be a valid and reliable tool for quantifying physical activity in adults.^{16–18} At the start of the data collection period, a research assistant visited each participant at home and measured his or her height (Seca Stadiometer, Hamburg, Germany) and weight (Seca Scales, Hamburg, Germany) with footwear and heavy clothing removed. An ActiGraph accelerometer was then attached to each subject's right hip via an adjustable elastic belt. At the same time subjects were given instructions on how to fit the device and a contact number if problems occurred. Each of the participants wore the ActiGraph during all waking hours for 7 consecutive days. The ActiGraph device was set to record at 1 minute epochs, with data for each day considered valid only if 10 or more hours of data were collected. Raw accelerometer counts were downloaded into the ActiLife computer software for determination of time spent in light 0–1951 counts (1.0–2.9 multiple of resting metabolic rate (METs)), moderate 1952–5724 counts (3.0–6.0 METs), vigorous ≥ 5725 counts (≥ 6.0 METs) and moderate-to-vigorous activity ≥ 1952 counts (≥ 3.0 METs). The physical activity cut-offs corresponding to these intensities were derived from the prediction equation of Freedson *et al* (1998),¹⁹ and are similar to the intensity categories used in other research studies.^{18 20 21} To calculate all activity that may be important for disease prevention, time spent in activity of a defined intensity (e.g. moderate or vigorous) was determined by summing the

individual 1 minute epochs in the day where the count met the criterion for that intensity.

In a counterbalanced manner, participants were required to complete both the NZPAQ-LF and the IPAQ-LF self-report questionnaires immediately after the 7 day accelerometer monitoring period. Data gained from the IPAQ-LF was coded and analysed using the recommended guidelines found on the IPAQ website (www.ipaq.ki.se). Using the IPAQ scoring system, the total number of days and minutes of physical activity were calculated for each participant in the areas of moderate and vigorous-intensity activity and total physical activity (moderate + vigorous-intensity). Each participant was also given a categorical score of “Low”, “Moderate” or “High” according to their level of activity as outlined in the IPAQ guidelines.

The NZPAQ-LF was administered by the researcher, who used the New Zealand Sports and Physical Activity Survey Showcards and Interviewer Guidelines found on the SPARC website (www.sparc.org.nz) to assist the participants to complete the survey. The minutes of moderate and vigorous activity were summed to give daily and weekly physical activity totals. Activity totals were then used to classify participants as either “Active” or “Inactive” according to their ability to meet the New Zealand physical activity guidelines for duration of physical activity (≥ 150 min/week) or duration and frequency of physical activity (≥ 30 min/day on ≥ 5 days per week).¹⁵

Statistical analysis

No editing of the data was performed except when input error was detected after checking against original records. All data sets obtained from the ActiGraph physical activity monitors and the two questionnaires were compiled in a Microsoft Excel spreadsheet and then transferred to the Statistical Analysis System (Version 8.2, SAS Institute, Cary NC) for further analysis. Means and standard deviations were calculated for minutes of moderate, vigorous and total (moderate and vigorous) physical activity per week using data obtained from each of the three instruments. Independent *t* tests were used to determine significant differences between groups. Spearman correlation coefficients were calculated to compare total weekly minutes of moderate to vigorous-intensity physical activity measured by the ActiGraph and self-reported on the NZPAQ-LF and IPAQ-LF. We used Cohen's²² guidelines for classifying the correlations (i.e. $r < 0.30$, small; $r = 0.31$ – 0.50 , moderate; $r > 0.50$, large). The level of disagreement in physical activity levels between the three collection methods was estimated by the Bland–Altman method,²³ in which differences between collection methods are plotted against their averages. Nominal variables representing the proportion of participants (and subgroups) meeting current physical activity guidelines were compared in SAS by categorical modelling using general linear modelling.

RESULTS

A total of 70 adults were surveyed during this study; however, only 64 data sets were included in the analysis, as six sets of data were found to be incomplete due to either the failure of participants to wear the ActiGraph physical activity monitor for the entire 7 day period or the failure of the participants to correctly complete either or both of the NZPAQ or the IPAQ. The participants involved in this study are a representative sample of the total population in terms of gender, but our sample contained about double the proportion of 18–35-year-olds that is present in the true population.²⁴ The characteristics of the research participants are displayed in table 1 below.

Table 1 Participant Characteristics

	N (% of sample)	Age (years)	Height (m)	Weight (kg)	BMI (kg/m ²)
Total	64 (100)	35.2 (14.4)	1.73 (0.09)	72.1 (13.6)	24.5 (4.1)
Age (years)					
18–35	33 (51)	22.3 (4.2)	1.72 (0.10)	67.4 (10.9)	23.5 (4.6)
36–50	19 (30)	45.3 (3.8)	1.71 (0.08)	75.8 (13.8)	24.9 (2.7)
51–65	12 (19)	54.9 (3.4)	1.78 (0.07)	79.3 (15.7)	26.5 (4.2)
Gender					
Male	28 (44)	34.8 (15.4)	1.79 (0.07)	66.9 (11.4)	23.6 (4.9)
Female	36 (56)	35.6 (13.7)	1.68 (0.07)	76.1 (13.9)	25.1 (3.3)

Data are presented as mean (SD).

Self-reported physical activity levels

In comparison with the ActiGraph, self-reported levels of moderate, vigorous and total (moderate and vigorous-intensity) physical activity were substantially overestimated in both the NZPAQ and the IPAQ questionnaires. Mean total (moderate + vigorous-intensity) physical activity levels measured via the two self-report questionnaires were approximately 165% higher than the ActiGraph-measured activity levels (table 2). The variance for the total time spent in moderate and vigorous-intensity physical activity was substantially lower for the ActiGraph than for the two self-report measures, indicating a greater spread of scores for the subjective questionnaires (table 2). Despite this overestimation in the mean activity levels, there were moderate to strong correlations between objectively measured and self-reported data (table 3).

Although the ActiGraph had lower overall variance than the two self-report measures, day-to-day variation in the moderate to vigorous-intensity ActiGraph data was reasonably high (26 minutes, 95% confidence limits 20–32 minutes) considering the New Zealand physical activity guideline cut-off (30 minutes).

Bland–Altman comparisons on the moderate–vigorous activity data indicated good agreement between the objectively measured ActiGraph and self-reported NZPAQ-LF data up to approximately 500 min/week. However, as physical activity levels increased, the NZPAQ-LF tended to overestimate moderate–vigorous physical activity (fig 1). Similarly, good agreement was found between the ActiGraph and self-reported IPAQ-LF data up to approximately 1000 min/week. As physical activity levels increased over 1000 min/week, the IPAQ-LF tended to overestimate moderate–vigorous physical activity (fig 2).

Data from the two self-report questionnaires showed high correlation with each other. Mean minutes of moderate-intensity ($r = 0.68$), moderate–vigorous-intensity ($r = 0.79$),

and vigorous-intensity ($r = 0.67$) physical activity from both instruments were significantly correlated and the Bland–Altman plot showed good levels of agreement between the two instruments throughout the physical activity range (fig 3).

Proportion of participants meeting physical activity guidelines

In comparison with the ActiGraph data, the two self-report questionnaires overestimated the proportion of the population meeting the duration and frequency physical activity guideline (table 4). Although there was closer agreement between the questionnaires and the ActiGraph data when it came to predicting the physical activity guideline based on duration only, the NZPAQ-LF results remained significantly higher than the ActiGraph results.

DISCUSSION

Both the NZPAQ-LF and the IPAQ-LF demonstrated acceptable levels of validity when compared with data gained from the ActiGraph physical activity monitor. Validity coefficients for total physical activity (moderate–vigorous intensity) ranged from 0.30 to 0.32 and are similar to what is typically reported ($r = 0.20$ to 0.40) for such questionnaires.^{25 26} Notwithstanding the reasonable validity coefficients, the objective data gained from the ActiGraph indicate that both the self-report instruments tend to overestimate minutes of total, moderate and vigorous physical activity. These findings are comparable with those of a number of other major validation studies conducted in New Zealand and worldwide, which show that, when compared with objective physical activity data gained from either accelerometry or heart rate monitoring, self-report questionnaires have acceptable validity, but typically overestimate total physical activity.^{12 13 15 27 28}

Although the overall validity for both questionnaires was reasonable, one must take care in using such instruments in

Table 2 Physical activity accumulated over 7 days as measured by ActiGraph MTI accelerometer, NZPAQ-LF and IPAQ-LF

	ActiGraph			NZPAQ			IPAQ		
	Mod	Vig	Total	Mod	Vig	Total	Mod	Vig	Total
Total	317 (256)	41 (70)	357 (283)	826* (998)	123* (230)	949* (1076)	665* (845)	276*,† (623)	940* (1152)
Age (years)									
18–35	289 (292)	41 (79)	330 (327)	890* (1031)	163 (298)	1053* (1179)	645 (813)	441*,† (808)	1086* (1318)
36–50	353 (204)	51 (53)	385 (211)	810 (1094)	90 (126)	900 (1077)	751 (1038)	134 (295)	885 (1105)
51–65	336 (233)	53 (72)	389 (267)	677 (787)	66 (95)	743 (780)	582 (615)	46 (71)	628 (612)
Gender									
Male	363 (346)	50 (64)	413 (377)	1003* (1212)	132 (224)	1135* (1313)	870* (1096)	413*,† (752)	1283* (1530)
Female	280 (150)	33 (74)	314 (173)	688* (785)	116 (238)	805* (840)	505 (546)	169 (484)	674 (645)

Mod; moderate-intensity, Vig; vigorous-intensity, Total; moderate and vigorous-intensity.

Data are in minutes, presented as mean (SD).

*significantly different from ActiGraph.

†significantly different from NZPAQ-LF.

Table 3 Agreement of objectively measured and self-reported physical activity duration

	ActiGraph vs NZPAQ			ActiGraph vs IPAQ		
	Mod	Vig	Mod-Vig	Mod	Vig	Mod-Vig
Total	0.30 (0.06–0.51)	0.37 (0.14–0.56)	0.32 (0.08–0.52)	0.19 (–0.06–0.42)	0.42 (0.19–0.60)	0.30 (0.06–0.51)
Age (years)						
18–35	0.42 (0.19–0.60)	0.46 (0.24–0.63)	0.41 (0.18–0.60)	0.09 (0.16–0.33)	0.36 (0.13–0.56)	0.33 (0.09–0.53)
36–50	0.44 (0.22–0.62)	0.30 (0.06–0.51)	0.46 (0.24–0.63)	0.32 (0.08–0.52)	0.61 (0.43–0.74)	0.36 (0.13–0.56)
51–65	0.08 (–0.17–0.32)	0.27 (0.03–0.48)	0.10 (–0.15–0.34)	0.37 (0.14–0.56)	0.46 (0.24–0.63)	0.36 (0.13–0.56)
Gender						
Male	0.38 (0.15–0.57)	0.34 (0.10–0.54)	0.37 (0.14–0.56)	0.25 (0.00–0.47)	0.25 (0.00–0.47)	0.29 (0.05–0.50)
Female	0.28 (0.04–0.49)	0.42 (0.19–0.60)	0.31 (0.07–0.52)	0.14 (–0.11–0.37)	0.54 (0.34–0.69)	0.30 (0.06–0.51)

Mod, moderate-intensity; Vig, vigorous-intensity; Mod-Vig, moderate to vigorous-intensity physical activity. Data are Spearman correlation coefficients (95% confidence limits).

subsets of the population. The considerably lower validity correlations between ActiGraph and NZPAQ-LF data in the 51–65-year-olds suggests that care may be needed in using the NZPAQ-LF with these subjects. Further research is needed to elucidate the reasons behind such low correlations.

Both the IPAC-LF and NZPAQ-LF produced substantially higher (approximately 165%) physical activity mean scores than the ActiGraph accelerometer. This lack of agreement is due to either overestimation by the self-report questionnaires or underestimation of the true physical activity levels by the ActiGraph device. Overestimation, which has been reported to be as high as 200–300%,^{29–30} can be due to a variety of factors. Self-report questionnaires are essentially dependent on the accuracy of participant recall, and for this reason a high error rate can exist. Furthermore, in today’s health and exercise conscious society individuals often overreport socially desirable behaviours such as physical activity and underreport socially undesirable behaviours such as inactive or sedentary behaviours, thus influencing the outcomes of self-reported research.³¹ Recent data suggests that social-desirability bias is associated with overreporting of physical activity duration by approximately 4–11 minutes per day over a 7 day period.³¹

Variability between measurement instruments may also be due to differences in the determination of activity levels. For example, at present there is disagreement about what ActiGraph cut-points should be used to determine moderate and vigorous-intensity activity. The current study used cut-points described by Freedson *et al* (1998); however, a more

recent study suggests slightly higher cut-points.³² Consequently, using different cut-points will alter the subsequent physical activity data and possibly agreement between measurement instruments.

On the other hand, the disagreement between the self-reported and objectively measured results in this study may be due to underestimation of activity levels by the ActiGraph accelerometer. It was interesting to find such high similarity between overall total physical activity scores between the IPAQ-LF and NZPAQ-LF instruments (less than 1%) in the present study. This agreement was also reflected in the Bland–Altman plot (fig 3) and the high correlation coefficient in the mean total physical activity minutes between these two independent questionnaires ($r = 0.79$), which may indicate that it was the ActiGraph underreporting rather than the self-report overestimating that caused the difference between the two types of instruments.

While a recent study found that the ActiGraph had a high level of interinstrument and intrainstrument reliability compared with other accelerometry-based activity monitors,³³ it also significantly underestimated energy expenditure when tested on individuals while they were participating in physical activity.³⁴ Evidence suggests that this underestimation is partly due to the inability of accelerometers to detect activity which does not involve movement in the vertical plane from the hip where the monitor is worn. This results in poor detection of such activities as cycling, rowing and upper body resistance training, and may have resulted in underestimation of activity

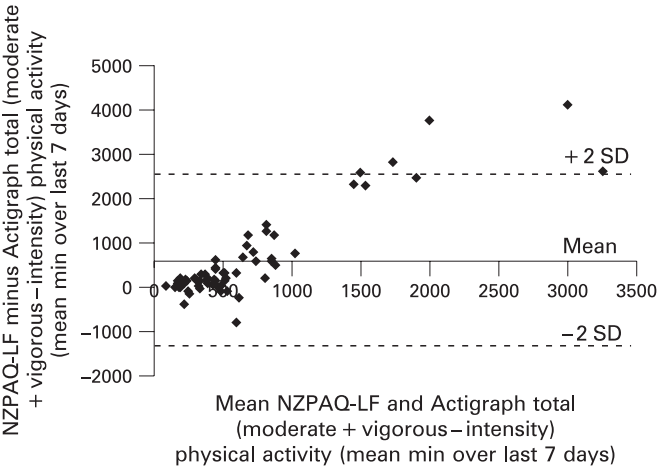


Figure 1 Agreement between self-reported (NZPAQ-LF) and objectively measured (ActiGraph Accelerometer) total (moderate + vigorous-intensity) physical activity duration.

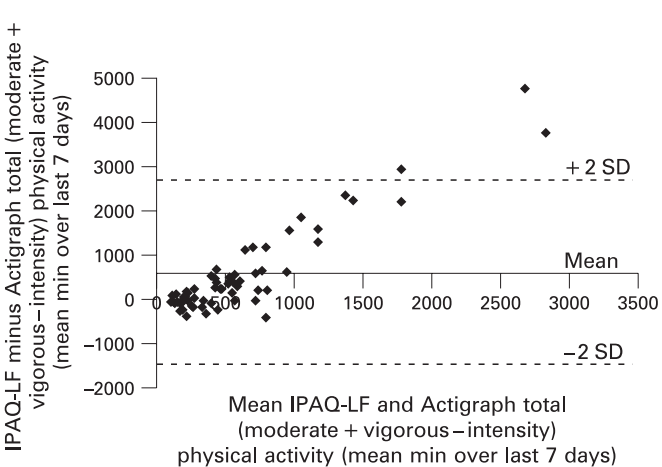


Figure 2 Agreement between self-reported (IPAQ-LF) and objectively measured (ActiGraph Accelerometer) total (moderate + vigorous-intensity) physical activity duration.

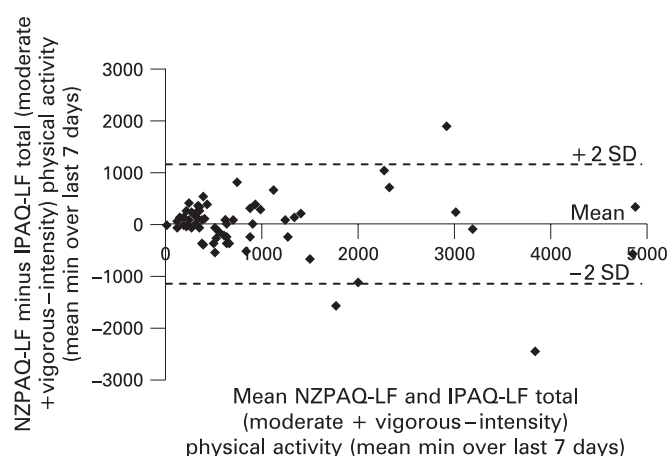


Figure 3 Agreement between self-reported NZPAQ-LF and IPAQ-LF total (moderate + vigorous-intensity) physical activity duration.

levels in the subjects of this study. On the other hand, it may be unrealistic to expect very high agreement between an instrument that has very strict physical activity cut-offs and the less precise self-report instruments. Recent data would tend to support this suggestion. In a unique study in which the authors used six different instruments to measure health-enhancing physical activity (two self-report and four objective), levels of activity showed high agreement within the different types of instrument (r = approximately 0.6 for self-reports and r = approximately 0.4 to 0.7 for objective measures) but poor agreement between instruments (r = approximately -0.1 to 0.3).³⁵ Overall physical activity levels calculated from the self-reports were approximately 330% higher than from the objectively measured instruments. These results suggest that the disagreement between instruments is probably caused by overreporting of the self-completed questionnaires rather than underreporting of the objective instruments, since it seems unlikely that all four objective instruments underreported activity levels to the same degree.

The Bland-Altman plots showing the level of agreement between the IPAQ-LF, NZPAQ-LF and ActiGraph data (figs 1 and 2) indicated much better agreement at lower levels of physical activity. This discrepancy may demonstrate a general inability of participants to differentiate between light-intensity and moderate to vigorous-intensity activity. The best example of the inability of participants to distinguish between the physical and physiological differences associated with the shift from light-intensity activity to moderate-intensity activity is

the large number of research participants who recorded workplace activity as equating to 8 h or 480 minutes of moderate activity per day. ActiGraph data showed that, although the majority of these participants took a large number of steps per day, suggesting that they were not employed in sedentary jobs, the activity involved in their jobs was not intense enough to be rated as moderate-intensity physical activity by the ActiGraph, thus leading to significant overestimation of daily physical activity. Moy *et al* (2008)¹⁵ suggested that clarification of the terminology used to define light, moderate and vigorous-intensity physical activity would help to minimise bias resulting from misinterpretation.

Clearly individuals vary their physical activity levels considerably throughout the week, as evidenced by the typical day-to-day variation in the ActiGraph-measured moderate-to-vigorous physical activity of 26 minutes. Such variation can cause differences in results depending on how physical activity is collected and interpreted. For example, the overall moderate to vigorous physical activity level of our subjects was on average 51 minutes per day (table 2); however, only 45% of these subjects reached the physical activity guidelines of 30 minutes per day on most days of the week (table 4). This discrepancy is due to a number of subjects being highly active on only a few days of the week. Such variation shows the limitation of using physical activity measurements that employ recording over 1 day rather than several days.

A substantial limitation of this study is the inability of the ActiGraph GT1M monitor to be worn during all waking hours. Unlike heart rate monitors, the ActiGraph is not waterproof, and therefore during this study participants had to remove the monitoring device when showering or participating in any aquatic activities such as swimming or kayaking. This means that the data collected by the ActiGraph may not be a true representation of activity undertaken during all waking hours over a 7 day period. However, research participants were instructed to remove the monitor only if absolutely necessary, in an attempt to keep physical activity data not recorded by the ActiGraph to a minimum. Including heart rate monitoring with accelerometry may be beneficial in future studies, as the continuous heart rate can provide useful information when no movement is occurring to the ActiGraph (e.g. arm movement or cycling) or when accelerometers have problems measuring true energy output (e.g. uphill walking).³⁶

A further limitation is that our study involved a relatively small sample of people from one location, which may limit the applicability of the results to the wider population. In addition, the cut-points for what constitutes moderate or vigorous-intensity

Table 4 Proportion of participants meeting New Zealand physical activity guideline cut-offs as measured by the ActiGraph, NZPAQ-LF and IPAQ-LF

	n	ActiGraph		NZPAQ		IPAQ	
		≥30 min most days	≥150 min over week	≥30 min most days	≥150 min over week	≥30 min most days	≥150 min over week
Total	64	45	83	78*	97*	75*	88
Age (years)							
18–35	33	36	79	85*	97*	79*	91
36–50	19	58	95	84	100	79	89
51–65	12	50	75	50	92	58	75
Gender							
Male	28	46	82	64	93	79*	89
Female	36	44	83	89*	100	72*	86

Data are percentages of participants in each subgroup meeting the New Zealand physical activity guideline cut-offs of ≥30 min most days (accumulate at least 30 min of physical activity on 5 out of 7 days of the week) and ≥150 min over week (accumulate at least 150 min of physical activity over the week).

*significantly different from ActiGraph at 0.05 level.

What is already known on this topic

Research investigating the validity of large-scale self-report physical activity questionnaires, typically used to measure the effectiveness of population-wide physical activity interventions in New Zealand, is scarce.

What this study adds

Validation of the longer formats of the New Zealand Physical Activity Questionnaire and the International Physical Activity Questionnaire against accelerometry (ActiGraph GT1M) in a New Zealand adult population revealed that such questionnaires typically overestimate physical activity.

activity are also possible limitations. Using alternative cut-points can change the amount of activity for each participant and therefore the overall results of the study.

CONCLUSION

The NZPAQ-LF and the IPAQ-LF demonstrated satisfactory levels of validity in this study of a randomised sample of adult New Zealanders, and results of the two self-report instruments are strongly correlated. Further work is required on validation with other objective measures of physical activity to ascertain the amount of overestimation in such questionnaires. However, since the validity of the NZPAQ-LF and the IPAQ-LF is similar to that of other self-report physical activity questionnaires, they can be considered acceptable instruments for measuring and assessing physical activity on a population scale.

Acknowledgements The Environment, Society and Design Division of Lincoln University funded a Summer Research Studentship for Rachel Boon to complete this project.

Competing interests None.

REFERENCES

1. Dunn AL, Marcus BH, Kampert JB, *et al.* Comparison of lifestyle and structured interventions to increase physical activity and cardiorespiratory fitness. *JAMA* 1999;**281**:327–34.
2. US Department of Health and Human Services. Physical activity for health: A report of the Surgeon General. Atlanta: Department of Health and Human Service, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, 1996.
3. Brownson RC, Boehmer TK, Luke DA. Declining rates of physical activity in the United States: What are the contributors? *Annu Rev Public Health* 2005;**26**:421–44.
4. Katzmarzyk PT, Gledhill N, Shephard RJ. The economic burden of physical inactivity in Canada. *Can Med Assoc J* 2000;**163**:1435–40.
5. Department of Health. Choosing Health? Choosing Activity: a consultation on how to increase physical activity. London: Department of Health, 2004.
6. Bauman A, Armstrong T, Davies J, *et al.* Trends in physical activity participation and the impact of integrated campaigns among Australian adults 1997–1999. *Aust N Z J Public Health* 2003;**27**:76–9.
7. Sport and Recreation New Zealand. SPARC Facts: Results of the New Zealand sport and physical activity surveys (1997–2001). Wellington: SPARC Policy Research Team, 2003.
8. Duncan JS, Schofield G, Duncan EK. Pedometer-determined physical activity and body composition in New Zealand children. *Med Sci Sports Exerc* 2006;**38**:1402–9.
9. Bassett DRJ. Validity and reliability issues in objective monitoring of physical activity. *Res Q Exerc Sport* 2000;**71**:30–6.
10. Schmidt MD, Freedson PS, Chasan-Taber L. Estimating physical activity using the CSA accelerometer and a physical activity log. *Med Sci Sports Exerc* 2003;**35**:1605–11.
11. Armstrong N, Bray S. Physical activity patterns defined by continuous heart rate monitoring. *Arch Dis Child* 1991;**66**:245–7.
12. Craig CL, Marshall AL, Sjostrom M, *et al.* International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;**35**:1381–95.
13. Mäder U, Martin BW, Schutz Y, *et al.* Validity of four short physical activity questionnaires in middle aged persons. *Med Sci Sports Exerc* 2006;**39**:1255–66.
14. Maddison R, Mhurchu CN, Jiang Y, *et al.* International Physical Activity Questionnaire (IPAQ) and New Zealand Physical Activity Questionnaire (NZPAQ): A doubly labelled water validation. *Int J Behav Nutr Phys Act* 2007;**4**:62.
15. Moy KL, Scragg RK, McLean G, *et al.* The New Zealand physical activity questionnaires: Validation by heart-rate monitoring in a multiethnic population. *J Phys Act Health* 2008;**5**:S45–S61.
16. Melanson EL Jr, Freedson PS. Validity of the Computer Science and Applications, Inc. (CSA) activity monitor. *Med Sci Sports Exerc* 1995;**27**:934–40.
17. Welk GJ, Schaben JA, Morrow JR Jr. Reliability of accelerometry-based activity monitors: a generalizability study. *Med Sci Sports Exerc* 2004;**36**:1637–45.
18. Brage S, Wedderkopp N, Franks PW, *et al.* Reexamination of validity and reliability of the CSA monitor in walking and running. *Med Sci Sports Exerc* 2003;**35**:1447–54.
19. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc* 1998;**30**:777–81.
20. Hendelman D, Miller K, Baggett C, *et al.* Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Med Sci Sports Exerc* 2000;**32**:S442–9.
21. Leenders NY, Nelson TE, Sherman WM. Ability of different physical activity monitors to detect movement during treadmill walking. *Int J Sports Med* 2003;**24**:43–50.
22. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. New Jersey: Lawrence Erlbaum, 1988.
23. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307–10.
24. Statistics New Zealand. Population Census Counts. <http://www.stats.govt.nz/people/population/populationcensuscounts.htm> (accessed 22 May 2008).
25. Bauman A, Merom D. Measurement and surveillance of physical activity in Australia: an introductory guide. *Aust Epidemiol* 2002;**9**:2–5.
26. de Courten M. Developing a simple global physical activity questionnaire for population studies. *Aust Epidemiol* 2002;**9**:6–9.
27. Brown W, Bauman A, Chey T, *et al.* Comparison of surveys used to measure physical activity. *Aust N Z J Public Health* 2004;**28**:128–34.
28. Hagstromer M, Oja P, Sjostrom M. The International Physical Activity Questionnaire (IPAQ): A study of concurrent and construct validity. *Public Health Nutr* 2006;**9**:755–62.
29. Macfarlane DJ, Lee CC, Ho EY, *et al.* Reliability and validity of the Chinese version of IPAQ (short, last 7 days). *J Sci Med Sport* 2007;**10**:45–51.
30. Rzewnicki R, Vanden Auweele Y, De Bourdeaudhuij I. Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. *Public Health Nutr* 2003;**6**:299–305.
31. Adams SW, Matthews CE, Ebbeling CB, *et al.* The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol* 2005;**161**:389–98.
32. Metzger JS, Catellier DJ, Evenson KR, *et al.* Patterns of objectively measured physical activity in the United States. *Med Sci Sports Exerc* 2008;**40**:630–8.
33. Eslinger DW, Tremblay MS. Technical reliability assessment of three accelerometer models in a mechanical setup. *Med Sci Sports Exerc* 2006;**38**:2173–81.
34. Leenders NY, Sherman WM, Nagaraja HN. Energy expenditure estimated by accelerometry and doubly labelled water: Do they agree? *Med Sci Sports Exerc* 2006;**38**:2165–72.
35. Macfarlane DJ, Lee CCY, Ho EYK, *et al.* Convergent validity of six methods to assess physical activity in daily life. *J Appl Physiol* 2006;**101**:1328–34.
36. Terrier P, Aminian K, Schutz Y. Can accelerometry accurately predict the energy cost of uphill/downhill walking? *Ergonomics* 2001;**44**:48–62.