Reliability and validity of a self-administered Arabic version of the Global Physical Activity Questionnaire (GPAQ-A)

Author

Doyle, Ciaran, Khan, Asaduzzaman, Burton, Nicola

Published

2019

Journal Title

Journal of Sports Medicine and Physical Fitness

Version

Accepted Manuscript (AM)

DOI

10.23736/S0022-4707.18.09186-7

Doyle, C.B, Khan, A., Burton, N.W. Reliability and validity of self-administered Arabic version of the Global Physical Activity Questionnaire (GPAQ-A). *Journal of Sports Medicine and Physical Fitness.* Accepted 9 October 2018.

Abstract

BACKGROUND: The primary objective of this study was to assess the test-retest reliability of the Arabic language version of the Global Physical Activity Questionnaire (GPAQ-A). Additionally, the study assessed criterion validity of the instrument.

METHODS: Emirati university students (n=93) were asked to complete the GPAQ-A on two occasions (7 days apart). A subset of 48 participants wore an accelerometer for 7 days after which they completed the GPAQ-A.

RESULTS: Test-retest reliability of the GPAQ-A was acceptable for moderate to vigorous physical activity (MVPA) ($\rho = 0.78$; 95% CI: 0.69 to 0.86), moderate physical activity (MPA) ($\rho = 0.73$; 95% CI: 0.61 to 0.82) and vigorous physical activity (VPA) ($\rho = 0.62$; 95% CI: 0.44 to 0.76). Test-retest reliability of the single item sedentary behaviour (SB) measure was less promising ($\rho = 0.44$; 95% CI: 0.22 to 0.64). Criterion validity of the MVPA measure was fair ($\rho = 0.23$; 95% CI: -0.13 to 0.57).

CONCLUSIONS: Test-retest reliability for the GPAQ-A was acceptable to assess physical activity (PA) but less promising for SB. Criterion validity of the PA measures was fair. Additional work is needed to develop robust measures of PA and SB in this population.

Key Words: exercise, psychometrics, sedentary behaviour, survey, Arabic

Introduction

Previous research has identified that regular physical activity (PA) participation is associated with a reduced risk of all-cause mortality, cardiovascular disease, type 2 diabetes, hypertension, breast cancer, colon cancer, gestational diabetes, ischemic heart disease, and ischemic stroke.[1] PA is also associated with a range of psychological health benefits including reduced risk of depression,[2, 3] anxiety [3] and stress [2, 3] as well as improved mood. [2] These physical and psychological benefits may be limited to specific domains of PA. For physical benefits, some research indicates that PA in the leisure, but not the work, domain is positively associated with cardiovascular health and inversely associated with mortality risk.[4, 5] For psychological benefits, PA performed in leisure and transport domains, but not in the work domain, is positively associated with mental health.[6]

Despite the benefits of regular PA participation, it is estimated that 23% of the world's adult population is insufficiently active.[7] Physical inactivity is a global pandemic causing a worldwide disease burden in morbidity and mortality.[8] In addition to disease burden, a review of economic burden conservatively estimated that illnesses associated with physical inactivity cost health care systems approximately $53.8 billion in 2013 and were responsible for 13.4 million disability adjusted life years worldwide.[9]

Inactivity is an emerging public health issue among the Emirati population. The discovery of oil in the 1960s has led to the rapid economic development, urbanisation and industrialisation of the United Arab Emirates (UAE). This has contributed to changes in lifestyle habits and health behaviours of Emirati people including less PA and increased sedentary behaviour (SB).[10-12] These lifestyle changes have resulted in significant public health implications with a rise in prevalence of inactivity-related conditions such as type 2 diabetes from 12.6% of population aged 20-79 in 2011 to 15.6% in 2017 amongst the residents

of the UAE.[13, 14] Furthermore, non-communicable diseases, many of which are inactivity related, are the cause of 65% of all deaths in the 30-70 year age group in the UAE.[15]

Given the importance of PA for physical and mental health, and the health and economic burden of inactivity, monitoring PA participation, and evaluating strategies to promote activity participation, is essential. Self-administered questionnaires are popular for large epidemiological studies as they are low cost, and offer greater speed and flexibility in data collection than objective measures. Furthermore, they can be easily administered using online methods which reduce participant burden. Questionnaires can also provide data on domain specific activities. However, ensuring acceptable psychometric properties of self-report questionnaires is essential.[16]

A systematic review of PA and SB research in the Arabian Gulf noted an urgent need for the use of standardised measures.[17] The Global Physical Activity Questionnaire – Arabic (GPAQ-A) is one of the most commonly utilised PA measures in the Arabian Gulf region;[17] but the test-retest reliability and criterion validity are yet to be adequately understood. Initial work assessing test-retest reliability and criterion validity of GPAQ-A with Saudi university students demonstrated moderate to strong test-retest correlation for MVPA and weak criterion validity with accelerometer derived MVPA.[18] However, this study only sampled males (n=67) and did not use all items in the questionnaire (6 work items were excluded). Furthermore, the participants were helped to complete the questionnaire by research assistants; this procedure may have created an external influence on participant responses. More work is needed therefore, to assess the psychometric properties of the GPAQ-A, with a self-administration procedure and including both female and male respondents.

The primary objective of the current study was to assess test-retest reliability of the 16 item GPAQ-A in a mixed gender sample of Arabic speakers using an online self-administered

version of the questionnaire. A secondary aim was to assess criterion validity of the questionnaire.

Materials and methods

This study was cleared in accordance with the ethical review guidelines and processes of The University of Queensland (HMS16/1905.R2). These guidelines are endorsed by the Human Experimentation Ethical Review Committee, and registered with the Australian Health Ethics Committee as complying with the National Statement.

Participants

All participants were Arabic speaking university students recruited from a Higher Education Institute in the Middle East. The university offers undergraduate programs and has separate campus for males and females. Study inclusion criteria were: male or female aged over 18 years, UAE national and native Arabic speaker. To assess test-retest reliability, participants were asked to complete the GPAQ-A on two occasions seven days apart. Participants were asked to create a unique identifier code (a three-digit code combined with day and month of birth) which allowed for matching of data across assessments. On completion of the first study, people were invited to participate in a second study assessing the criterion validity of the GPAQ-A. Participant engagement for both studies is outlined in Figure 1.

Test-Retest Reliability

Sample size for the reliability analyses was estimated using previously recommended procedures.[19] To achieve a power of 90% with a minimum acceptable reliability of 0.6 and a minimum desired reliability of 0.7; the required sample size for the 16-item questionnaire was

102. Assuming an initial response rate of 60%, and attrition of 25% between time 1 and time 2, 227 people were invited to the study at time 1. Students from 17 undergraduate classes (Business and Health majors) were invited to participate in the study via classroom visits. All participants were offered 2- hours credit as part of the university's student volunteer initiative (each student is required to complete 100 hours volunteering before completion of their degree).

Criterion Validity

All participants who had completed the test-retest reliability study were sent an email invitation to the study assessing the criterion validity of the GPAQ-A. All participants of the validity study were offered 14 hours' credit as part of the university's student volunteer initiative.

Materials

Global Physical Activity Questionnaire- Arabic (GPAQ-A)

GPAQ Version 2 consists of 16 items with a recall period of a typical week. The GPAQ is split into three domains: work, transport, and recreation. Vigorous PA (VPA), described in the questionnaire as "activities that require hard physical effort and cause large increases in breathing or heart rate", is assessed for domains of work (e.g., paid and unpaid work, study and/or household chores) and recreation (e.g., sports, fitness, and leisure). Each domain has three items measuring engagement (yes/no), frequency (days per week) and duration (hours and minutes per day). Moderate PA (MPA), described in the questionnaire as "activities that require moderate physical effort and cause small increases in breathing or heart rate", is assessed using work, recreation, and transport (the usual way you travel to and from places) domains. There is also a single item for SB, described as sitting or reclining at work, at home,

getting to and from places, or with friends, but not time spent sleeping and measuring time (hours and minutes) on a typical day. The Arabic version of the GPAQ was accessed from the World Health Organisation (WHO) website at http://www.who.int/ncds/surveillance/steps/resources/GPAQ_Analysis_Guide.pdf.[20]

Accelerometer measured time spent in PA and SB (criterion validity)

Participants were asked to wear an Actigraph GT3X accelerometer on a waist belt on the right hip during waking hours for seven consecutive days and for a minimum of 10 hours per day. They were also asked to record wear and non-wear times for each day in a log book. Participants were instructed to remove the device when sleeping or engaging in water-based activities.

The accelerometer was set to record in 10 second intervals which were aggregated into 60 second epochs. Periods of consecutive strings of zero count epochs lasting 60 minutes or more were considered as non-wear time. Epochs with more than 20,000 counts per minute were considered spurious[21] and removed from analysis. Interruption intervals of up to 2 consecutive epochs of <100 counts that appeared in the middle of long strings of zeros were removed.[22] Freedson cut off points (Sedentary < 100 counts per minute; Moderate: 1952 - 5724 counts per minute; Vigorous: 5725 - 9498 counts per minute)[23] were used when classifying the accelerometer data. These are the most commonly used cut off points in accelerometer research and have been used in previous studies evaluating criterion validity of PA questionnaires.[24]

Demographics

Participants also completed demographic information on age (years), gender, ability to manage on available income (impossible, difficult all of the time, difficult some of the time,

not too bad, easy), height and weight (used to derive body mass index), life satisfaction (rated from 1 to 10), and overall general health (excellent, very good, good, fair, poor). Some variables were categorized for descriptive purposes: age (18-20, 21-24, 25+ years) and life satisfaction (1-3 = low; 4-7 = middle; 7-10 = high).

Procedure

In the test-retest reliability study, the first author attended classes and discussed the study with students. Participants were then directed to the online survey link (which included a prerequisite online consent form) and completed the self-administered GPAQ-A on two occasions, seven days apart. They were also asked to provide demographic information at Time 1. In the validity study, a subset of interested people met with the first author, provided written consent and were briefed on how to wear the accelerometer and complete a wear time log. Participants were asked to wear the device for seven consecutive days after which they completed the GPAQ-A online.

Statistical Analysis

For test-retest reliability and criterion validity studies, the GPAQ-A data were screened and cleaned according to the procedures in the WHO STEPS Surveillance Manual.[25] A MPA variable was created by summing moderate work, transport, and moderate recreation domains. A VPA variable was created by summing vigorous work and vigorous recreation domains. The MVPA variable was created by summing MPA and VPA. In the criterion validity study, all cases where the participant wore the device for a minimum of ten hours per day for at least five days and including one weekend day were included in the analyses. Accelerometer data were aggregated and averaged to determine time (mins/day) spent in MVPA and SB. All data were

manually entered into IBM SPSS Statistics for Windows Version 24. Ten percent of the data entered were checked for accuracy and no errors were found.

Test-retest reliability of GPAQ-A was conducted for each of the self-reported MPA, VPA, and MVPA as well as SB in minutes per day. In the validity study comparisons were made between: 1) total self-reported MVPA and average accelerometer recorded MVPA, and 2) total self-reported SB and average accelerometer recorded SB.

For test-retest reliability, Bland-Altman plots were performed to assess mean bias and 95% limits of agreement between Time 1 and Time 2 data for PA (MPA, VPA, and MVPA) and SB.[26] For criterion validity, Bland-Altman plots were performed to mean bias and 95% limits of agreement between GPAQ-A and accelerometer data for MVPA and SB.[26] As the PA and SB data were not normally distributed, relationships were assessed using Spearman's rho coefficient (ρ), as has been done in other similar studies.[18, 27-31] Spearman's rho coefficient was interpreted as: 0-0.20 poor correlation, 0.21-0.40 fair correlation, 0.41-0.60 moderate/acceptable correlation, 0.61-0.80 substantial correlation, 0.81-1.0 near perfect correlation.[32] Differences in median minutes per day for Time 1 and Time 2 were assessed using Wilcoxon-signed rank tests.

Results

Participant characteristics

The participant engagement is detailed in Figure 1. For test-retest reliability, n=227 people were invited to participate in the study. Of these, 93 (41%) completed the survey on both occasions (Time 1 and 2) and passed data screening. The majority of the sample was female (59%) and the mean age of participants was 21.27 (SD 2.32) years. Of those invited to participate, 41 (18%) provided consent but did not complete the study or were excluded during

data screening. Of these, the majority were female (63%) and the mean age was 22.0 (SD 2.90) years.

For the criterion validity study, an invitation was sent to all who had completed the test-retest reliability study (n=93). Of these, n=48 consented to participate and n=43 (response rate 46%) completed the study and passed data screening (i.e. at least 5 days including 1 weekend day, and at least 10hrs/day wear time). The majority of the sample in the validity study was female (60.5%) and the mean age of participants was 19.30 (SD 1.87) years. Additional characteristics of the participants in both studies are described in Table 1.

Test-retest reliability

Results of the test-retest reliability analyses are shown in Table 2. For MVPA, Bland-Altman plots (see Figure 2A) showed that the difference between the measures was 5.88 mins/day. However, limits of agreement were wide and ranged from -95.9 to +107.67 mins/day. There was a substantial correlation between Time 1 and Time 2 measures ($\rho = 0.78$; 95% CI: .69 to .86). The median of reported MVPA times was 77 minutes per day for Time 1 and 86 mins/day for Time 2. This difference was not statistically significant ($p = .316$).

For VPA, Bland-Altman plots (see Figure 2B) showed a difference of 0.34 mins/day. Limits of agreement were wide and ranged from -58 to +58 mins/day. There was a substantial correlation between Time 1 and Time 2 ($\rho = 0.62$; 95% CI: 0.44 to 0.76). The VPA medians (Time 1 = 9 mins/day; Time 2 = 17 mins/day) were not significantly different ($p = .836$).

For MPA, Bland-Altman plots (see Figure 2C) showed a difference of 5.54 mins/day. Limits of agreement were wide and ranged from -90 to +101 mins/day. There was a substantial correlation between Time 1 and Time 2 ($\rho = 0.73$; 95% CI: 0.61 to 0.82). The medians (Time 1 = 51 mins/day; Time 2 = 57 mins/day) were not significantly different ($p = .308$).

For SB, Bland-Altman plots (see Figure 2D) showed a difference of 32.58 mins/day. Limits of agreement were wide and ranged from -487 to +552 mins/day. There was a moderate

correlation between Time 1 and Time 2 (ρ = 0.44; 95% CI: 0.22 to 0.64). The SB medians (Time 1 = 300 mins/day; Time 2 = 300 mins/day) were not significantly different (p = .139).

Criterion Validity

Results of the criterion validity analyses are shown in Table 3. Comparing self-reported MVPA and objectively measured MVPA, Bland-Altman plots (see Figure 3A) showed that the difference between the two measures was 61.1 mins/day. Limits of agreement were wide and ranged from -139 to +262 mins/day. There was a fair correlation between GPAQ-A measured and accelerometer measured MVPA (ρ = 0.23; 95% CI: -0.13 to 0.57). GPAQ-A measured median of MVPA was 43 mins/day, while the median of MVPA measured by the accelerometer was 20 mins/day. This difference was statistically significant (p< = 0.001).

Comparing self-reported SB and objectively measured SB, BA plots (see Figure 3B) showed that the difference between the two measures was 99 mins/day. Limits of agreement were wide and ranged from -539 to +341 mins/day. There was a poor correlation between GPAQ-A measured and accelerometer measured SB (ρ = -0.02; CI: -0.35 to 0.33). Median times for SB was 360 mins/day for GPAQ-A data and was 492 mins/day for accelerometer data. This was a statistically significant difference (p = 0.008).

Discussion

The results demonstrate that the GPAQ-A showed acceptable test-retest reliability for MVPA, VPA, and MPA across one week. Reliability coefficients for PA in our study are comparable with studies that used a similar test-retest period. Our MVPA reliability coefficient was comparable with interview and self-administered English language versions in Singapore [33] and a French version tested in France.[34] Our reliability coefficients for MPA are higher than reported in those studies; however, our reliability coefficients for VPA are lower.

An interesting point to note is that the MPA measure had a higher reliability coefficient than that of the VPA measure. This is in contrast to other studies that have assessed test-retest reliability of different language versions of GPAQ,[30, 33, 34] and is difficult to interpret. Previous research has reported that vigorous intensity activities may be easier to recall as they are more purposive and structured than moderate intensity activities.[35-37] The corollary to this is that moderate intensity activities (e.g., brisk walking) may be more difficult to recall as they are more varied, and may be accumulated over the entire day. Further exploration of how activities of different intensities are perceived and recalled may be required in this population.

Test-retest reliability of the single item SB measure showed a poor correlation. Our results for SB are less favourable than the French language version, which was the only other study that used a similar test-retest interval,[34] and less favourable than other studies that used different test-retest intervals assessing versions of the GPAQ in English in America, Arabic in Saudi Arabia, and French in France.[18, 30, 34] It is possible that estimating SB using a single item measure may be difficult for the respondent as SB may occur across multiple domains. Furthermore, SB is an unstructured behaviour and consequently may be difficult to recall over the course of a day.

The current study demonstrates that the GPAQ-A showed fair criterion validity for MVPA when compared with objectively measured data. Our results are generally comparable with some similar GPAQ studies from Saudi Arabia (Arabic), Malaysia (Bahasa Melayu), the United States (English, Spanish), and Singapore (English).[18, 27, 28, 30, 33] Results for the criterion validity of the SB item showed no relationship between self-report and objectively measured data. Studies from Saudi Arabia (Arabic), the United States (English), and Northern Ireland (English) have also reported poor criterion validity for the GPAQ SB measure[18, 30, 38] whilst studies from the United States (Spanish), Bangladesh (Bangla), Switzerland (German), and France (French) have reported moderate to fair validity.[28, 29, 31, 34] Developing valid and reliable

measures of SB is important for epidemiological and intervention studies.[39] The results from the current study suggest that more work developing valid Arabic language self-report measures of SB is required. Recent research using domain specific measurement of SB has had promising results.[40, 41]

A strength of the current study is the utilisation of accelerometers as the criterion for objective PA measure. Accelerometers are more robust in the measurement of PA intensities than some measures used in other studies e.g. pedometers.[42] Conversely, using Actigraph accelerometers as the referent measure for SB may have been a limitation as past research has identified that this type of accelerometer may not be the most sensitive objective measure of SB.[39] The study used convenience sampling of university students and hence the results may not be applicable to the wider population. More research is needed therefore to assess the psychometrics of the questionnaire across age groups.

Conclusions

In conclusion, this study adds new information on the test-retest reliability and criterion validity of the GPAQ-A in Arabic speaking populations. The GPAQ-A, as a PA measure, shows some promise as a reliable population surveillance tool in the UAE. The GPAQ-A does not provide a valid measure of SB and more work is needed to develop and evaluate Arabic language measures of SB, in particular that provide domain specific data on sedentariness.

References

1. Warburton DER, Bredin SSD. Health benefits of physical activity: a systematic review of current systematic reviews. Current Opinion in Cardiology. 2017;32(5):541-56.
2. Penedo JF, Dahn RJ. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. Current Opinion in Psychiatry. 2005;18(2):189-93.
3. Rebar AL, Stanton R, Geard D, Short C, Duncan MJ, Vandelanotte C. A Meta-Meta-Analysis of the effect of physical activity on depression and anxiety in non-clinical adult populations. Health Psychology Review. 2015:1-78.
4. Li J, Loerbroks A, Angerer P. Physical activity and risk of cardiovascular disease: what does the new epidemiological evidence show? Current Opinion in Cardiology. 2013;28(5):575-83.
5. Richard A, Martin B, Wanner M, Eichholzer M, Rohrmann S. Effects of leisure-time and occupational physical activity on total mortality risk in NHANES III according to sex, ethnicity, central obesity, and age. J Phys Act Health. 2015;12(2):184-92.
6. White RL, Babic MJ, Parker PD, Lubans DR, Astell-Burt T, Lonsdale C. Domain-specific physical activity and mental health: A Meta-analysis. American Journal of Preventive Medicine. 2017;52(5):653-666.
7. World Health Organisation. Prevalence of insufficient physical activity. 2010 [cited 2018 22nd February]; Available from: http://www.who.int/gho/ncd/risk_factors/physical_activity_text/en/.
8. Kohl HW, Craig CL, Lambert EV, Inoue S, Alkandari JR, Leetongin G, et al. The pandemic of physical inactivity: global action for public health. The Lancet. 2012;380(9838):294-305.
9. Ding D, Lawson KD, Kolbe-Alexander TL, Finkelstein EA, Katzmarzyk PT, van Mechelen W, et al. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. The Lancet. 2016; 388:1311-1324.
10. Dhaheri AS, Mohamad MN, Jarrar AH, Ohuma EO, Ismail LC, Meqbaali FT, et al. A cross-sectional study of the prevalence of metabolic syndrome among young female Emirati adults. PloS one. 2016;11(7).
11. Ng SW, Zaghloul S, Ali HI, Harrison G, Popkin BM. The prevalence and trends of overweight, obesity and nutrition-related non-communicable diseases in the Arabian Gulf States. Obesity Reviews. 2011;12(1):1-13.
12. Sabri S, Bener A, Eapen V, Abu Zeid MSO, Al-Mazrouei AM, Singh J. Some risk factors for hypertension in the United Arab Emirates. Eastern Mediterranean Health Journal. 2004;10(4-5):610-9.
13. International Diabetes Federation. IDF Diabetes Atlas (5th Ed.). 2011.
14. International Diabetes Federation. IDF Diabetes Atlas (8th Ed.). 2017.
15. World Health Organisation. World Health Organization - Noncommunicable Diseases (NCD) Country Profiles Geneva: Switzerland: 2014.
16. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): Nine country reliability and validity study. Journal of Physical Activity and Health. 2009;6(6):790-804.
17. Mabry R, Koohsari MJ, Bull F, Owen N. A systematic review of physical activity and sedentary behaviour research in the oil-producing countries of the Arabian Peninsula. BMC Public Health. 2016; 16:1003.
18. Alkahtani SA. Convergent validity: agreement between accelerometry and the Global Physical Activity Questionnaire in college-age Saudi men. BMC Research Notes. 2016;9(1):436.

19. Nunnally JC, Bernstein IH. Psychometric theory. 3rd ed. New York: McGraw-Hill; 1994.
20. World Health Organisation. Global Physical Activity Surveillance. [cited 2018 1st February 2018]; Available from: http://www.who.int/ncds/surveillance/steps/GPAQ/en/.
21. Masse LC, Fuemmeler BF, Anderson CB, Matthews CE, Trost SG, Catellier DJ, et al. Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. Medicine and Science in Sports and Exercise. 2005;37(11 Suppl):S544-54.
22. Evenson KR, Terry JW. Assessment of differing definitions of accelerometer nonwear time. Research Quarterly for Exercise and Sport. 2009;80(2):355-62.
23. Freedson PS, Melanson E, Sirard J. Calibration of the computer science and applications, Inc. accelerometer. Medicine and Science in Sports and Exercise. 1998;30(5):777-81.
24. Craig LC, Marshall LA, Sjöström EM, Bauman LA, Booth EM, Ainsworth FB, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Medicine and Science in Sports and Exercise. 2003;35(8):1381-95.
25. World Health Organisation. World Health Organisation: WHO STEPS Surveillance Manual: the WHO STEPwise Approach to Chronic Disease Risk Factor Surveillance. Geneva, Switzerland;2005.
26. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307-10.
27. Soo KL, Wan Abdul Manan WM, Wan Suriati WN. The Bahasa Melayu version of the Global Physical Activity Questionnaire: reliability and validity study in Malaysia. Asia-Pacific Journal of Public Health. 2015;27(2):184-193.
28. Hoos T, Espinoza N, Marshall S, Arredondo EM. Validity of the Global Physical Activity Questionnaire (GPAQ) in adult Latinas. Journal of Physical Activity and Health. 2012;9(5):698-705.
29. Mumu SJ, Ali L, Barnett A, Merom D. Validity of the Global Physical Activity Questionnaire (GPAQ) in Bangladesh. BMC Public Health. 2017;17(1):650.
30. Herrmann SD, Heumann KJ, Der Ananian CA, Ainsworth BE. Validity and reliability of the Global Physical Activity Questionnaire (GPAQ). Measurement in Physical Education and Exercise Science. 2013;17(3):221-35.
31. Wanner M, Hartmann C, Pestoni G, Martin BW, Siegrist M, Martin-Diener E. Validation of the Global Physical Activity Questionnaire for self-administration in a European context. BMJ Open Sport and Exercise Medicine. 2017;3(1).
32. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.
33. Chu AH, Ng SH, Koh D, Muller-Riemenschneider F. Reliability and validity of the self- and interviewer-administered versions of the Global Physical Activity Questionnaire (GPAQ). PloS one. 2015;10(9).
34. Rivière F, Widad FZ, Speyer E, Erpelding ML, Escalon H, Vuillemin A. Reliability and validity of the French version of the Global Physical Activity Questionnaire. Journal of Sport and Health Science. 2016;7:339-345.
35. Bassett DR, Jr. Validity and reliability issues in objective monitoring of physical activity. Research Quarterly for Exercise and Sport. 2000;71(2):30-6.
36. Washburn RA, Heath GW, Jackson AW. Reliability and validity issues concerning large-scale surveillance of physical activity. Research Quarterly for Exercise and Sport. 2000;71(sup2):104-13.
37. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nutrition. 2006;9(6):755-62.

38. Cleland CL, Hunter RF, Kee F, Cupples ME, Sallis JF, Tully MA. Validity of the Global Physical Activity Questionnaire (GPAQ) in assessing levels and change in moderate-vigorous physical activity and sedentary behaviour. BMC Public Health. 2014;14(1).
39. Atkin AJ, Gorely T, Clemes SA, Yates T, Edwardson C, Brage S, et al. Methods of measurement in epidemiology: Sedentary behaviour. International Journal of Epidemiology. 2012;41(5):1460-71.
40. Marshall LA, Miller DY, Burton WN, Brown JW. Measuring total and domain-specific sitting: A study of reliability and validity. Medicine and Science in Sports and Exercise. 2010;42(6):1094-102.
41. Chau JY, van Der Ploeg HP, Dunn S, Kurko J, Bauman AE. A tool for measuring workers sitting time by domain: the Workforce Sitting Questionnaire. British Journal of Sports Medicine. 2011;45(15):1216.
42. Corder K, Brage S, Ekelund U. Accelerometers and pedometers: methodology and clinical application. Current Opinion in Clinical Nutrition & Metabolic Care. 2007;10(5):597-603.

Table 1. Characteristics of participants

	All		Women		Men	
	n	%	n	%	n	%
	93	100.0	55	59.1	38	40.9
Age group (years)						
18-20	42	45.2	25	45.5	17	44.7
21-24	45	48.4	27	49.1	18	47.4
25-32	6	6.4	3	5.4	3	7.9
Body mass index (kg/ m^2)						
Underweight or normal weight (<25)	61	65.6	40	72.7	21	55.3
Overweight (25-30)	14	15.1	9	16.4	5	13.2
Obese (>30)	18	19.3	6	10.9	12	31.5
General health						
Excellent or very good	74	79.6	46	83.6	28	73.7
Good	14	15.1	7	12.7	7	18.4
Fair or poor	5	5.3	2	3.7	3	7.9
Ability to manage on available income						
Easy or not too bad	63	67.7	39	70.9	24	63.2
Difficult some of the time	28	30.1	14	25.5	14	36.8
Difficult all of time or impossible	2	2.2	2	3.5	0	0.0
Life satisfaction						
High	72	77.4	43	78.2	29	76.3
Moderate	15	16.1	10	18.2	5	13.2
Low	6	6.5	2	3.6	4	10.5

Table 2. Test-retest reliability of PA and SB (mins/day) (n=93)

	Median (mins/day)	IQR	p-value*	Spearman rho	95% CI
MVPA (n=93)					
Time 1	77.14	18-154	p=.316	.78	.69 to .86
Time 2	85.71	39-136			
Moderate Physical Activity (n=93)					
Time 1	51.43	14-114	p=.308	.73	.61 to .82
Time 2	57.14	18-111			
Vigorous Physical Activity (n=93)					
Time 1	8.57	0-43	p=.836	.62	.44 to .76
Time 2	17.14	0-43			
Sedentary Behaviour (n=86)					
Time 1	300	230-533	p=.139	.44	.22 to .64
Time 2	300	180-458			

*derived from median tests

Table 3. Spearman correlation between GPAQ-A and accelerometer based estimates of PA (mins/day) (n=43)

	Median (mins/day)	IQR	p-value*	Spearman rho	CI
Moderate to Vigorous Physical Activity (n=43)					
GPAQ-A	43	17-129	P<.0001	.23	: -.13 to .57
Accelerometer	20	14-34			
Sedentary Behaviour (n=43)					
GPAQ-A	360	240-600	p=.008	-.02	-.35 to .33
Accelerometer	492	431-563			

*derived from median tests

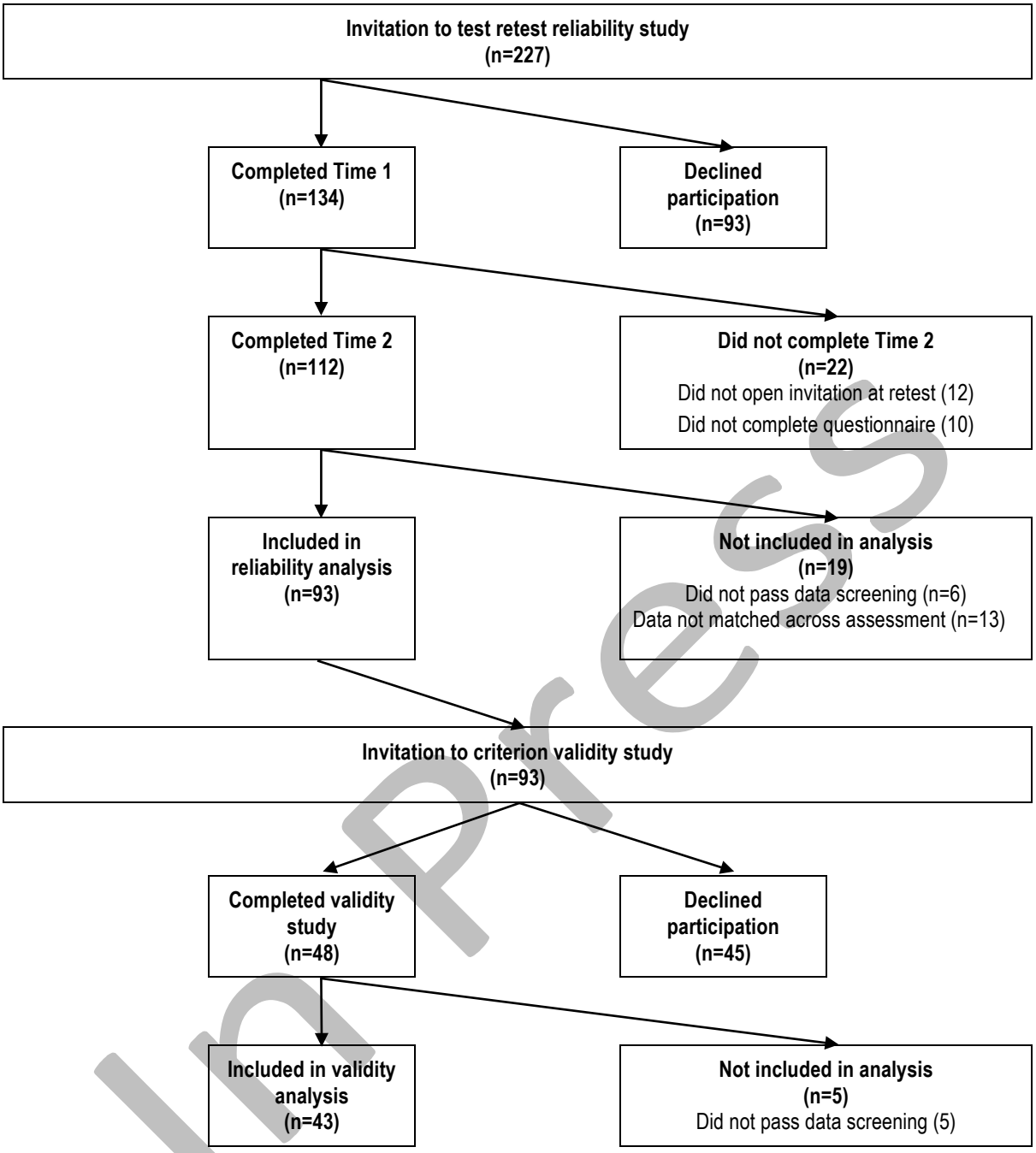
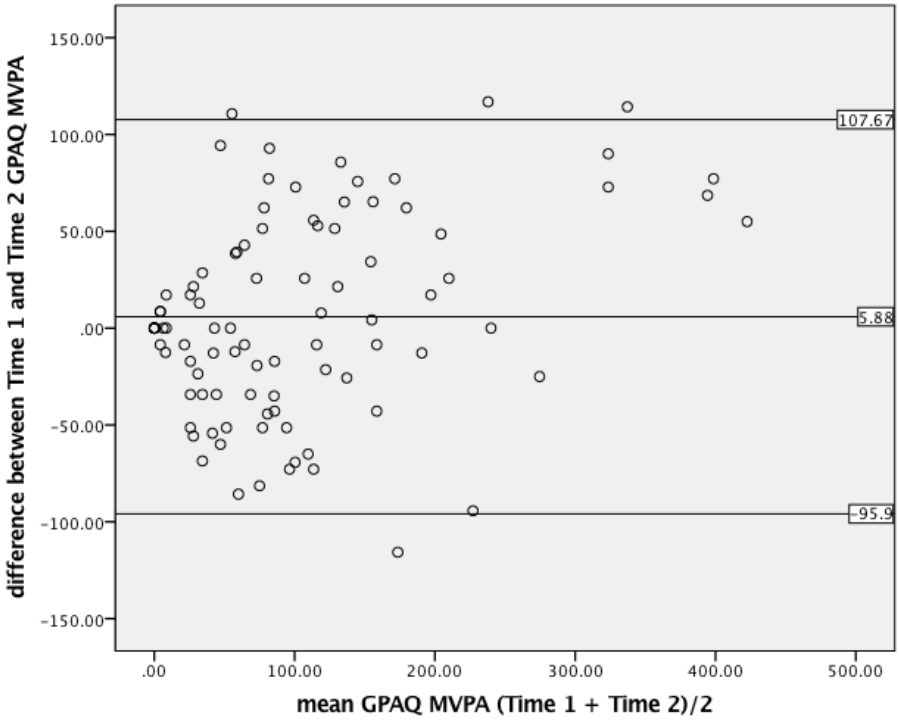
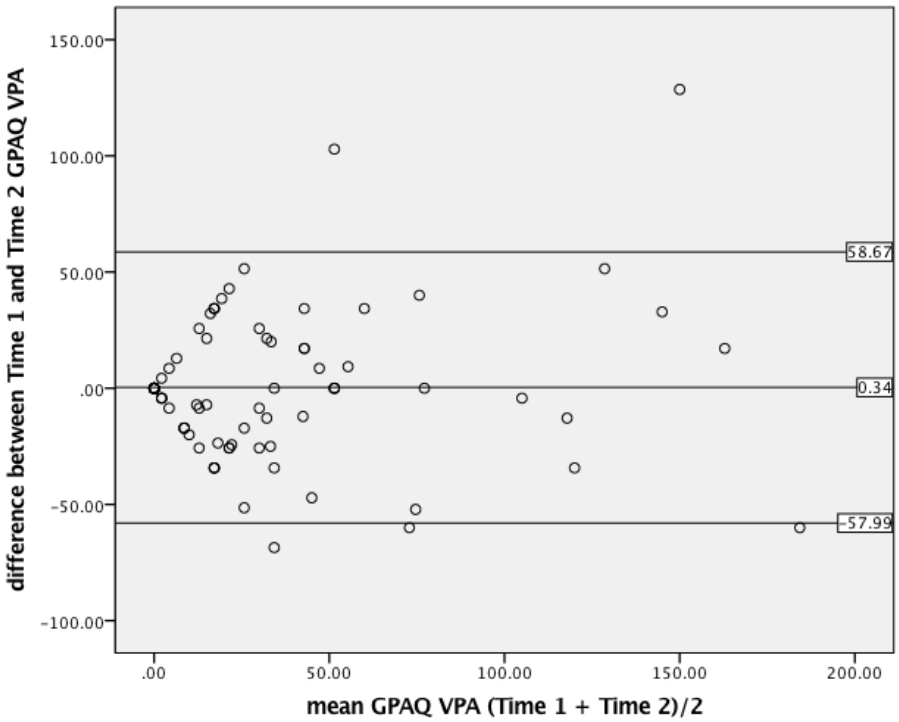


Fig 1. Flowchart of participants' engagement

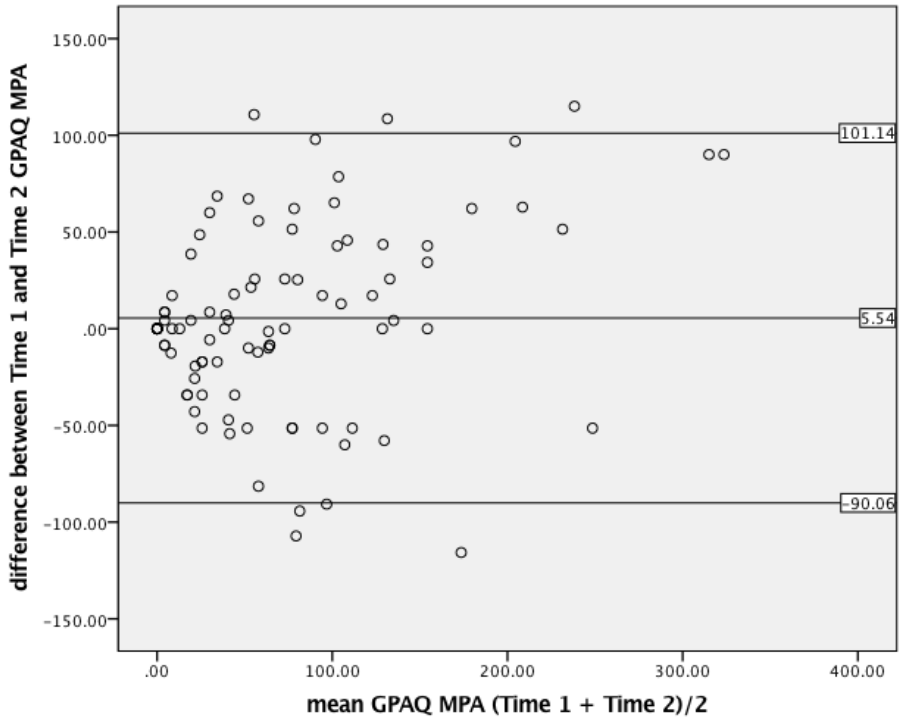
A


difference between Time 1 and Time 2 GPAQ MVPA
150.00
100.00
50.00
.00
-50.00
-100.00
-150.00
107.67
5.88
-95.9
.00
100.00
200.00
300.00
400.00
500.00
mean GPAQ MVPA (Time 1 + Time 2)/2

B


difference between Time 1 and Time 2 GPAQ VPA
150.00
100.00
50.00
.00
-50.00
-100.00
58.67
0.34
-57.99
.00
50.00
100.00
150.00
200.00
mean GPAQ VPA (Time 1 + Time 2)/2

C



D

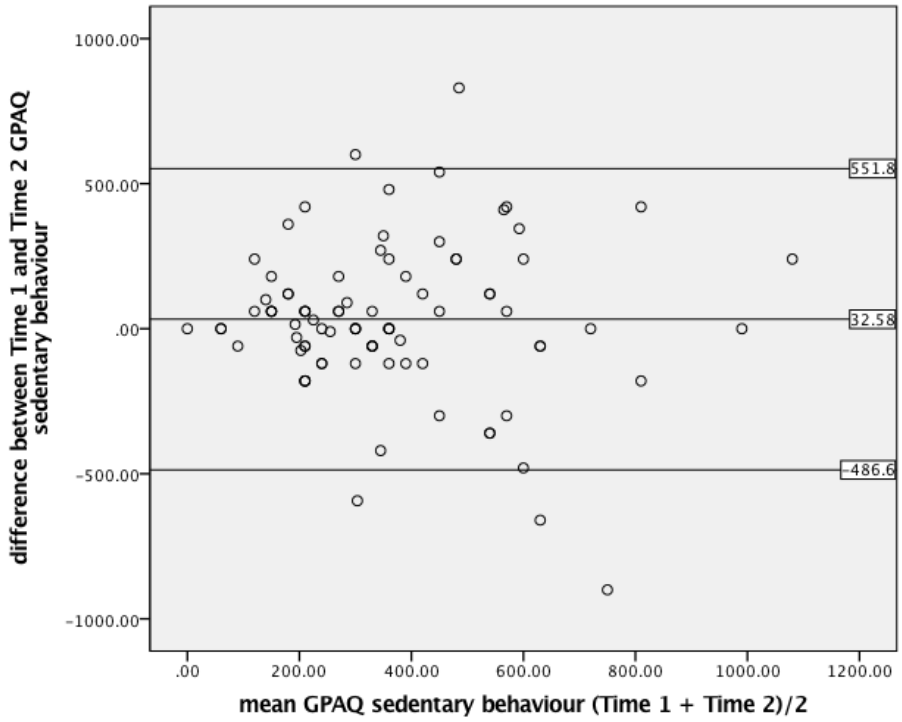
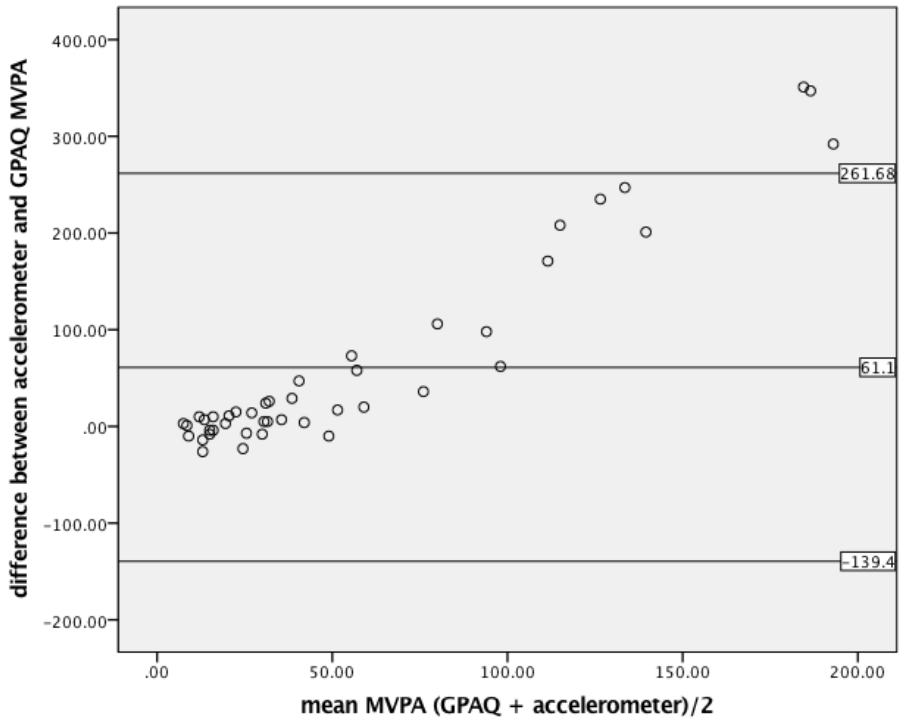


FIGURE 2: Bland Altman plot for the difference between (a) GPAQ MVPA mins/day at Time 1 and Time 2; (b) GPAQ VPA mins/day at Time 1 and Time 2; (c) GPAQ MPA mins/day at Time 1 and Time 2; (d) GPAQ SB mins/day at Time 1 and Time 2

A



B

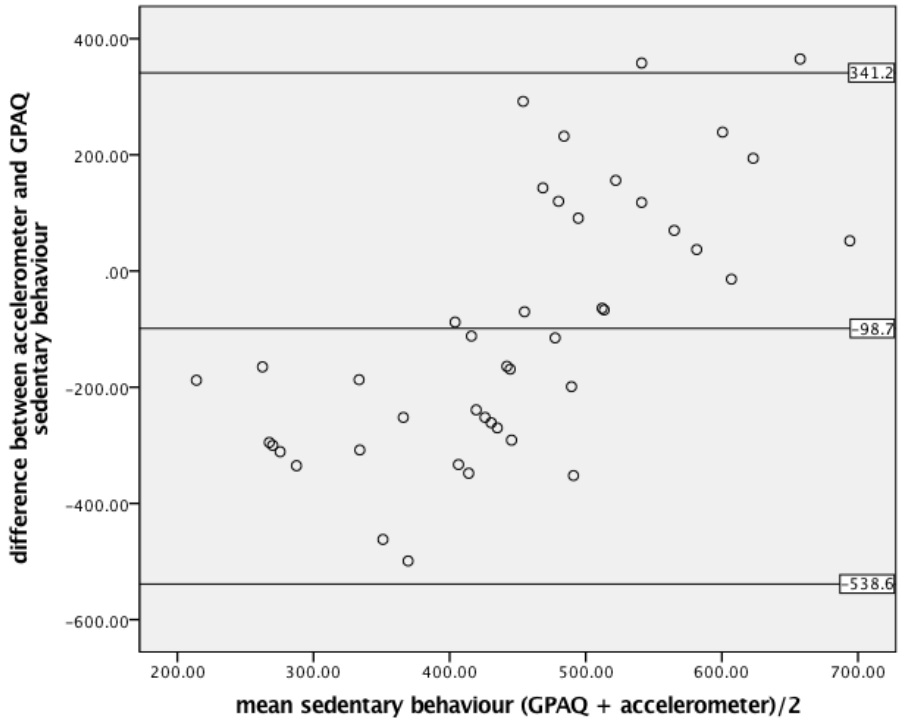


FIGURE 3: Bland Altman plot for the difference between (a) GPAQ MVPA mins/day and accelerometer measured MVPA mins/day; (b) GPAQ SB mins/day and accelerometer measured SB mins/day