The validity of the International Physical Activity Questionnaire (IPAQ) for adults with progressive muscle diseases

Sarah F. Roberts-Lewis, Claire M. White, Mark Ashworth and Michael R. Rose

School of Population Health and Environmental Sciences, King's College London, London, UK; Department of Neurology, King's College Hospital, London, UK

Abstract

Purpose: Measuring the physical activity of adults with progressive muscle diseases is important to inform clinical practice, for activity recommendations and for outcomes meaningful to participants in clinical trials. Despite its wide use, the measurement properties of the International Physical Activity Questionnaire (IPAQ) have not been established in a muscle disease population.

Materials and methods: The sample of 103 adults with progressive muscle diseases included independently mobile participants and wheelchair users. Their home-based activity measured by the IPAQ was compared to simultaneous weeks of accelerometer activity data collected remotely in a longitudinal, measure evaluation study. Validity, reliability, and responsiveness were evaluated for the IPAQ alone, and for the IPAQ used in conjunction with a smart activity monitor.

Results: The IPAQ did not demonstrate satisfactory criterion validity, reliability or responsiveness and it systematically overestimated moderate and vigorous physical activity time by 161 minutes per week. Measurement properties of the IPAQ were improved when it was used in combination with a smart activity monitor.

Conclusions: The IPAQ did not have satisfactory measurement properties compared to accelerometry in adults with progressive muscle disease. Combining self-report and objective activity measures might improve the accuracy of physical activity assessment in this and other comparable populations.

Introduction

Physical activity is defined as "any bodily movement produced by skeletal muscles that requires energy expenditure"; insufficient physical activity has been linked to poor health outcomes [1]. Measuring the physical activity of adults with progressive muscle diseases is important to inform clinical practice, activity recommendations and for outcomes, meaningful to participants, in clinical trials [2]. In muscle disease, activity has been linked to improved muscle strength, cardiorespiratory fitness, fatigue management, quality of life and protection against comorbidities [3–8].

However, concerns remain over potential activity-related exacerbations of muscle disease [4,9]. World Health Organisation (WHO) activity recommendations seem generally applicable and safe [1] for many other neuromuscular diseases. For example, those originating from neuromuscular junction pathology, such as Myasthenia Gravis [10], or nervous system pathologies, such as peripheral neuropathies [11,12] and Spinal and Bulbar Muscular Atrophy [12]. However, there is less evidence regarding safe activity dosage for adults with muscle disease [3]. The reciprocal relationship between activity and muscular gene expression [13] means that, in muscle disease, greater understanding and precision is required for measurement, and evidence-based prescription, of physical activity throughout disease progression [3,14].

Progressive muscle diseases, including muscular dystrophies, myopathies and inclusion body myositis, are characterised by pathological changes in skeletal muscles [15,16], resulting in a clinical course of slowly progressive weakness [17,18]. Disease symptoms are variable, for some there is a mobility decline that may require wheelchair use [19]. This functional heterogeneity makes activity measurement challenging. Multiple approaches have been used but there is insufficient evidence to recommend any activity measurement tool for use with adults who have progressive muscle diseases [20,21].
Despite not being designed originally for use in small samples for clinical studies [22], IPAQ is more widely used in neuromuscular disease research than objective measures or questionnaires tailored for disabled populations [21], such as the Physical Activity Scale for Individuals with Physical Disabilities [23]. Disability and wheelchair-specific activity questionnaires have received criticisms relating to the questionable accuracy of self-report [20] and may be suited to activity measurement only for those at later stages of disease progression. Continuity of activity measurement is preferable using a single measure inclusive of early and late stages of disease progression.

The International Physical Activity Questionnaire (IPAQ) was developed to standardise population-level activity surveillance worldwide [24,25]. Modified versions of the IPAQ have also been designed to evaluate people with functional limitations [26,27]. It is an easy to administer, self-report, 7-day recall questionnaire (or interview). It collects information about time spent in vigorous, moderate, walking, and sedentary activities. The overall score estimates metabolic expenditure and was designed to categorise people into low, moderate, or high activity. Early studies reported satisfactory reliability and validity in the general population [25,28,29]. However, subsequent systematic reviews indicated the IPAQ considerably overestimated activity and had only low to moderate validity when compared to objective measures [30–32]. IPAQ overestimation might be attributed to social desirability and recall biases [33] and systematic reviews in neuromuscular diseases have also questioned the accuracy of self-report physical activity measures [20,21].

However, if the IPAQ was valid for ambulant and non-ambulant adults with progressive muscle diseases, it would provide a suitable activity measurement tool to chart physical activity throughout the stages of muscle disease progression, confirm the accuracy of existing physical activity research in muscle disease, and allow comparisons with studies in the general population. To date, the IPAQ (modified or otherwise) has never been validated for people with progressive muscle diseases.

The objective of this study was to test the validity, reliability, and responsiveness of the IPAQ in adults with progressive muscle diseases. The secondary objective was to discover whether the measurement properties of the self-report questionnaire were improved by adjunctive smart activity monitoring. The hypothesis was that wearing a smart activity monitor would improve the accuracy of self-reported IPAQ physical activity data by reducing overestimation.

Materials and methods

Study design

Ethical approval for this longitudinal, measure evaluation study was granted by King’s College London Research Ethics Committee (LRS-18/19-10909). Recruitment started in April 2019. The setting was home-based as the study was conducted remotely. Baseline cross-sectional IPAQ and accelerometer data were collected from April to October 2019 and at longitudinal follow-up from February to August 2020. They were compared to evaluate validity, reliability, and responsiveness of the questionnaire. Measurement properties for the IPAQ alone were compared to those for IPAQ data collected after a week of adjunctive smart activity monitoring.

Participants

People with progressive muscle diseases were invited to participate via national muscle disease registries (The John Walton Muscular Dystrophy Research Centre, Newcastle University) and the charity, Muscular Dystrophy UK (https://www.musculardystrophyuk.org/) advertised via their website and newsletter. Participants were included if they were UK resident adults with a confirmed diagnosis of Inclusion Body Myositis, Myotonic Dystrophy or Muscular Dystrophy (including Facioscapulohumeral Dystrophy, Limb Girdle Muscular Dystrophy, Dysferlinopathy, Dystrophinopathy (including manifesting female carriers) or specific congenital myopathies lasting into adulthood). Participants were excluded if they were aged <18 years, cognitively impaired, unable to wear an accelerometer or did not have a confirmed diagnosis of muscle disease. After volunteers made contact with the research team, screening was carried out by telephone appointment. A physiotherapist (SRL) clinically assessed eligibility criteria and gathered supporting evidence (including diagnosis confirmation and cognitive history). Any eligibility discrepancies were resolved on a case-by-case basis by a consultant neurologist (MR). For eligible participants, informed consent and data were collected by email and post.

Outcomes

The primary outcome was concurrent validity of the IPAQ overall score correlated with overall activity intensity in mean accelerations per minute measured by accelerometer. Secondary outcomes included convergent, divergent, and discriminative validity according to demographic characteristics, test–retest reliability over 2 weeks and responsiveness to change from baseline to follow-up. Measurement error of IPAQ activity minutes compared to weekly moderate and vigorous minutes measured by accelerometer was also tested. Each measurement property was also evaluated for the IPAQ plus adjunctive smart activity monitoring.

Measures

The self-report, short-form IPAQ [25], consists of three questions about days and time of vigorous, moderate, and walking activity in bouts of ≥10 min, with one question about daily sedentary time in the preceding seven days. In this study, the question descriptors were modified to incorporate the range of functional ability in the sample; changes were similar to previous modified versions [26,27]. Modifications included extra vigorous and moderate activity examples, including wheelchair activities. The walking question was modified to include “time spent walking, self-propelled wheeling or equivalent light activities.” “Sedentary” time was changed to “inactive” time. The questionnaire was preceded by completion instructions and three questions about the types of vigorous, moderate, and light activity undertaken [34]. It was followed by three questions about bedtime, waketime and sleep hours. Weekly activity times were calculated from days and minutes of vigorous, moderate, and walking/light activity. Total score in metabolic equivalents (METS) per week was calculated by multiplying weekly active time by intensity specific metabolic values as per IPAQ scoring instructions [25].

Adjunctive smart activity monitoring was by Fitbit Inspire HR (Fitbit Inc., San Francisco, CA). It is a tri-axial accelerometer and continuous optical heart rate monitor. It syncs with the Fitbit smart phone app and yields physical activity metrics including daily active minutes, metabolic expenditure, steps, and sleep. In the adjunctive study, activity monitoring was used for a week preceding each IPAQ completion. Participants regularly synced the Fitbit with their smart phone app. Thus, participants had seen Fitbit activity metrics before completing the questionnaire at the end of the week.

The comparator measure was a GENEActiv tri-accelerometer (ActivInsights Ltd., Kimbolton, UK). It has been validated for...
wrist-wear [35] and ankle-wear in adults with progressive muscle
diseases [36]. Participants wore it on their non-dominant wrists
unless ankle-wear was indicated because of work or using
crutches (which can interfere with wrist-worn activity monitors).
The sampling frequency was 10 Hz. Data were processed using
the GGIR package in R (version 3.6.0) [37] in 1-min epochs of
milli-gravitational units (milli-g) with gravitational correction [38].
Weekly overall activity intensity was calculated by mean accelera-
tions (milli-g) per minute. Minutes of sleep, inactivity, light acti-
ity, and bouts of ≥10 min of moderate and vigorous activity were
calculated using the following cut-points: light ≥30 milli-g/min,
moderate ≥100 milli-g/min and vigorous activity ≥400/milli-g/
min [39].

The moderate to vigorous physical activity (MVPA) threshold of
100 milli-g/min has been widely used for healthy adults and peo-
ples with long-term conditions [40,41]. Lower thresholds were con-
sidered for this cohort, whose muscle weakness might mean
greater exertion produces less activity compared to healthy con-
trols [36]. However, muscle strength is not a direct determinant of
limb acceleration [36] and a <100 milli-g/min threshold would
have been misleading for more active individuals, including
wheelchair users, whose activity and exertion are equivalent to
people with other long-term conditions [42].

Procedure
Before baseline, demographic information was collected; these
data underpinned evaluation of IPAQ construct validity (discrim-
native, convergent, and divergent). Demographics included con-
dition, age, gender, anthropometrics, handedness, employment
status, mobility, quality of life and self-perception. Quality of life
was measured using the Individualised Neurological Quality of
Life (INQoL), a 45-item questionnaire designed and validated for
people with neurological diseases [43]. Items are Likert rated 0–6,
yielding an indexed score of 0–100 from best to worst quality of
life. Subscales include activities, independence, relationships, emo-
tions, and appearance. Self-perceived activity level was measured
using the Physical Self-Description Questionnaire activity subscale
(PSDQ-activity) [44]. The 40-item, Likert-rated PSDQ was validated
for disabled adults with neuromuscular diseases [45]. The PSDQ-
activity subscale of four items yields a mean score of 1–6 from
least to most activity.

At baseline weeks 1 and 2, test–retest reliability was tested by
repeated administration of the IPAQ and testing for functional sta-
bility each week using the Health Assessment Questionnaire
(HAQ) [46]. The 24-item HAQ has been validated in electronic and
paper format [47]. Activities of daily living are rated 0–3 according
to difficulty, yielding an indexed total score of 0–3 from least to
most disabled. At baseline week 3, concurrent validity and meas-
urement error were evaluated by comparison of IPAQ data with
GENEActiv accelerometer data. The GENEActiv was posted to par-
ticipants, who wore it for 7 days, only removing it to wash, then
returned it by post. The third baseline week was used to ensure
that accelerometry did not interfere with reliability testing of the
questionnaire in baseline weeks 1 and 2.

At follow up, participants were sent the accelerometer to wear
for another week and questionnaires were repeated. Responsiveness was examined by comparing changes in IPAQ and
accelerometry over time.

In the adjunctive study, 3 months after baseline, participants
were invited to wear a smart activity monitor and complete the
IPAQ again. Those who opted in were sent a wrist-worn Fitbit.
Participants used the Fitbit for 2 weeks at time points 1 and 2 and
during the follow up week. They completed the IPAQ and HAQ
each week to test adjunctive IPAQ test–retest reliability.
Participants who wore their Fitbit at follow up provided data to
test adjunctive IPAQ concurrent validity (compared to GENEActiv
accelerometry) and construct validity (relationships with demo-
graphics re-tested at follow up) (see Figure 1).

Statistical analysis
A sample size of 100 (plus 10% to allow for attrition) was planned
based on recommendations for evaluating measurement proper-
ties [48] and a power calculation [49]. Statistical analyses were
carried out using R (version 3.6.0). Alpha was set at 0.05. Data
were tested for normality using Shapiro–Wilk [50] and nonpara-
metric equivalents were used when data were not normally dist-
buted. MVPA was summed moderate and vigorous activity
time recorded using IPAQ/GENEActiv. Concurrent validity
between IPAQ and GENEActiv data was tested by Spearman’s
 correlation; rho ≥0.70 was considered satisfactory [48]. However,
lower correlations of 0.09–0.39 [31], have been reported and
≥0.50 is sometimes used as a satisfactory correlation cut-point
[31]. Based on a predicted correlation of 0.50–0.70, the study
had 99–100% power [49]. Convergent and divergent validity
were tested by correlation with related constructs (PSDQ-activity,
INQoL, and HAQ) and unrelated constructs (height, handedness,
and age) respectively. Satisfactory construct validity cut-points
were ≥0.30 and ≤0.20, respectively [48]. Discriminative validity
tested the difference in IPAQ score between wheelchair users
and ambulant participants by the Mann–Whitney U test.
Measurement error of activity time reported in the IPAQ was
examined by absolute measurement error and Bland–Altman
plots. Reliability was tested by intra-class correlation coefficient
(ICC) between repeated IPAQ scores one week apart. An ICC of
≥0.75 was considered satisfactory [51]. Responsiveness of the
IPAQ to change overtime was tested by area under the curve
(AUC) analysis of receiver operating characteristic curves for
GENEActiv change. An AUC of ≥0.70 was considered satisfactory
[48]. Questionnaires were scored using available items only;
questionnaires with >10% of items missing were excluded from
analyses. For accelerometry, missing data of ≤10 min were
included in daily means, but days with <2.3 h monitored were
excluded from analyses. Participants with missing datasets or
lost to follow up were excluded from analyses.

Results
Figure 1 summarises recruitment. One hundred and three com-
pleted baseline data; 100 returned follow up data. There was a
mean of 9.4 months between baseline and follow up.
Demographics at baseline are summarised in Table 1.

The concurrent validity between IPAQ score and GENEActiv
overall intensity did not meet acceptability criteria, although there
was a significant moderate correlation (rho = 0.49, p < 0.0001,
N = 103). Convergent validity of IPAQ score was demonstrated by
significant low to moderate correlations with the HAQ Disability
Index and PSDQ-Activity subscale (rho = −0.35 to 0.42).
Divergent validity of IPAQ score was demonstrated by lack of sig-
ificant correlation with height, handedness, employment status
and age (see Table 1). Discriminative validity was demonstrated
by significant differences between wheelchair users and inde-
pendently mobile participants for activity measured by IPAQ and
GENEActiv (see Tables 1 and 2). Test–retest reliability of the IPAQ
was just below the acceptability threshold (ICC = 0.73, confidence
interval $= 0.66–0.81$). There was no significant change in HAQ between reliability testing weeks. Responsiveness of the IPAQ to activity change over time was not satisfactory (AUC $= 0.64$). Using the IPAQ with the smart activity monitor (Fitbit) improved the validity and reliability of the questionnaire. However, criterion validity was still unsatisfactory (see Table 2).

The validity and test–retest reliability for activity time derivations of the IPAQ, including MVPA, light activity, inactivity, and sleep time were similarly unsatisfactory to that of IPAQ total score. However, like IPAQ total score, they showed slightly improved measurement properties for the IPAQ plus Fitbit compared to IPAQ alone (see Table 3).

Measurement error was high for IPAQ derived MVPA time, compared to the GENEActiv accelerometer, as shown by the Bland–Altman plot (see Figure 2(a)). There were wide limits of agreement, high proportional error (0.86) and considerable absolute and systematic errors indicating overestimation of MVPA using IPAQ alone (172 and 161 min per week, respectively). Those who reported fitness or transportation activities tended to overestimate less than respondents who reported only domestic activities. However, the types of activity reported were predominantly domestic (e.g., shopping, housework, self-care, house maintenance and caring responsibilities). Self-reported MVPA time showed the greatest overestimation error and inactivity time the greatest underestimation error. Whereas light activity and sleep time reporting had more random error (see Figure 2(a,c,e,g)).

IPAQ plus Fitbit MVPA showed the greatest improvement in self-report accuracy compared to IPAQ alone. MVPA measurement error reduced by 135 min per week and there was a 34% reduction in proportional bias (see Figure 2(b)). The effect was similar, but smaller, for light activity (see Figure 2(d)). However, there was little change in measurement error of inactivity or sleep time reporting between IPAQ alone and IPAQ plus Fitbit (see Figure 2(f,h)).
Table 1. Demographics and baseline information (N = 103).

| Characteristic | All | Independent | Assisted | Wheelchair users |
|---------------|-----|-------------|----------|------------------|
| Sample size   | 103 | 45          | 22       | 36               |
| Diagnoses     | 33 LGMD, 25 FSHD, 16 IBM, 11 MD, 10 CM | 11 LGMD, 16 FSHD, 5 IBM, 7 MD, 7 CM | 7 LGMD, 4 FSHD, 3 IBM, 2 MD, 10 CM | 15 LGMD, 5 FSHD, 8 IBM, 4 MD, 10 CM |
| Gender        | 48 Male, 55 Female | 19 Male, 26 Female | 12 Male, 10 Female | 17 Male, 19 Female |
| Employment status | 40FT, 17PT | 26FT, 6PT | 6FT, 6PT | 8FT, 5PT |
| Handedness    | 93 Right, 10 Left | 42 Right, 3 Left | 18 Right, 4 Left | 33 Right, 3 Left |
| Mobility      | 36 WC, 22 AM, 45 IM | 45 IM | 22AM | 36 WC |
| Age (years)   | mean (SD) | 47.8 (15.6) | 43.0 (14.7) | 49.4 (16.1) |
| Height (cm)   | mean (SD) | 170.3 (10.3) | 170.7 (11.1) | 171.8 (8.6) |
| Weight (kg)   | mean (SD) | 74.9 (17.6) | 70.3 (13.7) | 79.6 (19.1) |
| Body Mass Index (mean (SD)) | 25.8 (6.0) | 24.2 (4.8) | 26.8 (6.3) | 26.1 (6.1) |
| PSDQ-activity (median (IQR)) | 42.2 (36.7–49.4) | 37.8 (26.7–41.1) | 48.9 (44.8–52.1) | 48.9 (45.3–56.0) |
| INQoL (median (IQR)) | 46 NW | 12 NW | 12 NW | 22 NW |
| GenoActiv MVPA mins/week | 6.2 (10.8) | 12.3 (13.5) | 7.7 (16.0) | 1.7 (4.6) |
| IPAQ total score | 2604 (1317–4814) | 4484 (2920–6523) | 1985 (852–3432) | 1628 (679–2678) |
| IPAQ MVPA minutes | 107 (15–235) | 210 (120–360) | 50 (12–128) | 25 (0–105) |
| Accelerometer (mean (SD)) | Mean accelerations | 27.0 (9.1) | 34.1 (6.5) | 25.7 (7.4) |

All participants had at least five days of usable accelerometer data at baseline and follow up. All participants, except seven, wore the accelerometer on their wrists. Outputs did not differ significantly by position, so all accelerometer data were analysed together. Most participants completed electronic questionnaires. Only five at baseline and one at follow up, completed paper questionnaires. SD: standard deviation; IQR: interquartile range; MVPA: moderate and vigorous activity; HAQ (DI): Health Assessment Questionnaire (disability index); INQoL: Individualized Neurological Quality of Life questionnaire; LGMD: Limb Girdle Muscular Dystrophy; FSHD: facioscapulohumeral dystrophy; IBM: inclusion body myositis; DM: Myotonic Dystrophy; CM: congenital myopathies; NW: not working or retired; WC: wheelchair users or bedbound; AM: assisted mobility, ambulant with a walking aid or occasional wheelchair use outdoors for long distances; IM: independent mobility.

Table 2. Validity and reliability statistics for IPAQ total score.

| Measurement property | IPAQ (N = 103) | IPAQ plus Fitbit (N = 90) |
|----------------------|----------------|---------------------------|
| Concordant validity with IPAQ score | Statistic | p Value | Statistic | p Value |
| GENoActiv | rho = 0.49 | <0.0001 | rho = 0.57 | <0.0001 |
| Convergent validity with IPAQ score | PSQD-Activity subscale | rho = 0.42 | <0.0001 | rho = 0.57 | <0.0001 |
| HAQ Disability Index | rho = -0.35 | 0.0002 | rho = -0.37 | <0.0001 |
| INQoL-Independence | rho = -0.29 | 0.0032 | rho = -0.39 | <0.0001 |
| Weight | rho = -0.19 | 0.0438 | rho = -0.20 | 0.0321 |
| Divergent validity with IPAQ score | Handedness | rho = -0.13 | 0.1971 | rho = -0.03 | 0.7694 |
| Height | rho = -0.13 | 0.2075 | rho = -0.08 | 0.4449 |
| Employment status | rho = 0.10 | 0.3011 | rho = 0.09 | 0.2671 |
| Age | rho = 0.02 | 0.8445 | rho = 0.00 | 0.9686 |
| Gender | rho = 0.19 | 0.0615 | rho = 0.14 | 0.1591 |
| BMI | rho = -0.18 | 0.0743 | rho = -0.18 | 0.0622 |

| Discriminant validity between extreme groups (wheelchair users and independently mobile) | IPAQ overall score | U = 1316 | <0.0001 | U = 1444 | <0.0001 |
| GENoActiv overall intensity (milli-g/min) | U = 1589 | <0.0001 | U = 1616 | <0.0001 |
| Reliability of IPAQ score (between weeks 1 and 2 and timepoints 1 and 2, respectively) | Test-re-test | ICC = 0.73 | <0.0001 | ICC = 0.79 | <0.0001 |

IPAQ: International Physical Activity Questionnaire; HAQ: Health Assessment Questionnaire; INQoL: Individualized Neurological Quality of Life questionnaire; PSQD: Physical Self-Description Questionnaire; ICC: intra-class correlation coefficient.

Discussion

The IPAQ did not have satisfactory validity compared to accelerometry for the assessment of physical activity in adults with progressive muscle disease. Despite satisfactory construct validity (convergent, divergent, and discriminative), the IPAQ did not demonstrate satisfactory criterion validity, reliability, or responsiveness. IPAQ measurement error was unacceptably high with a strong tendency for overestimation of activity. However, measurement properties and reporting accuracy of the IPAQ were slightly...
improved when it was used in combination with a smart activity monitor.

Our findings are similar to other studies which have reported unsatisfactory validity and reliability of the IPAQ in healthy adults and clinical populations [53]. Conversely, some earlier studies reported IPAQ had satisfactory reliability and validity, however, they used lower acceptability thresholds [51]. Our findings of significant overestimations of IPAQ MVPA time are concordant with measurement error findings from other studies of self-reported physical activity measures [31,52,53]. Measurement error might be explained by recall and social desirability bias (which leads people to overestimate self-reported activity) [54].

Some researchers have suggested ways to improve the accuracy of self-report physical activity measurement [55]. It is possible that self-report bias might be tempered by adjunctive objective monitoring. This might explain the improved IPAQ plus Fitbit measurement properties detected in this study. However, some experts recommend preferential reliance on objective activity measures, only using self-report tools, like the IPAQ, alone, if objective measurement is unfeasible [52,53,56]. The IPAQ alone is unsuitable for clinical studies of adults with progressive muscle diseases, where activity represents a health outcome indicative of disease progression. Similarly, the IPAQ alone is unsuitable when the precise quantification of physical activity is crucial. For example, informing muscle disease activity prescription and guidelines, because inflated activity targets might lead to activity-related disease exacerbation [9] or activity disengagement. Insufficient activity might increase the detrimental health sequelae associated with sedentarism [57].

Self-report measures, like the IPAQ, are useful in population surveillance studies where activity categorisation, rather than exact activity quantification, is desirable [22,52]. Also, as additional activity outcomes to help interpret objective data and provide information about activity type and perceived intensity. Indeed, self-reported physical activity data provides insight into perceived activity intensity [30]. Self-reported MVPA overestimation, compared to accelerometry, might be attributable to the increased energy cost of physical activity experienced by some people with progressive muscle diseases [58,59]. Accelerometry is unable to differentiate the level of effort required for limb acceleration and some experts recommend combined activity measurement approaches encompassing movement and associated exertion [60].

In this study, the excessive activity overestimation of IPAQ alone might be explained by social desirability bias, recall bias and accelerometer insensitivity to exertion. Combined IPAQ plus adjunctive smart activity monitoring potentially represents a more accurate activity assessment, eliminating some self-report overestimation biases whilst also accounting for perceived energy cost of activity.

The improved accuracy of activity time reporting using the IPAQ plus smart activity monitoring might be accounted for because the monitoring app dashboard displayed active minutes. Thus, immediate availability of activity time information from smart activity monitoring might have facilitated more accurate MVPA (and light) activity self-report. However, self-report of inactivity time was consistently underestimated using IPAQ alone and unchanged with adjunctive smart activity monitoring. The monitoring app dashboard did not display inactivity data. This is a possible explanation for why smart activity monitoring did not improve awareness of inactivity. Thus, for assessment or interventions targeting inactivity, clinician-facilitated, or inactivity-focused, self-monitoring might be required to improve inactivity awareness and self-report.

The key strengths of this study were the thorough examination of multiple measurement properties and the remote design, allowing the inclusion of a wide range of participants from across England, Ireland, Scotland and Wales with varied functional ability and activity levels. Limitations include that the standard accelerometry cut points, used to differentiate vigorous, moderate, and light activity, might have been too severe for this population which included some highly disabled individuals. The measurement properties of the IPAQ may have been impacted by modifications to questionnaire descriptors or the administration method, because IPAQ was originally designed for telephone use which allows real-time response guidance (potentially reducing self-report errors) [22]. Responsiveness might have been deflated by the lack of clinically important change in activity of the whole sample at follow up after less than a year. However, responsiveness might also have been inflated by the reduced overestimation linked to Fitbit use at follow up. The study would have been strengthened by a larger sample size allowing sub-group analyses of measurement properties by muscle disease diagnosis.

In conclusion, the IPAQ, compared to accelerometry, was not sufficiently valid, reliable, or responsive to measure physical activity in adults with progressive muscle diseases. Adjunctive objective activity monitoring improved the measurement properties and accuracy of IPAQ self-reported physical activity. For accurate quantification of physical activity, we recommend using objective or combined approaches to physical activity measurement in preference to self-report alone. Given our diverse sample, this recommendation is likely to be generalisable to activity measurement in other long-term conditions and, even, healthy adults.
Acknowledgements

The authors thank The John Walton Muscular Dystrophy Research Centre at Newcastle University and the Muscular Dystrophy UK charity for their help with advertising the study.

Disclosure statement

The authors report no conflicts of interest.

Figure 2. Bland–Altman plots of IPAQ measured MVPA time compared to GENEActiv MVPA time (minutes per week).

Systematic error = 161 minutes per week overestimation
Standard deviation = 321 minutes
Proportional bias = 0.86
Absolute measurement error = 172 minutes per week

Systematic error = 46 minutes per week overestimation
Standard deviation = 367 minutes
Proportional bias = 0.00
Absolute measurement error = 279 minutes per week

Systematic error = 1475 minutes (24.6 hours) per week underestimation
Standard deviation = 1918 minutes (32.0 hours)
Proportional bias = 0.36
Absolute measurement error = 2030 minutes (33.8 hours) per week

Systematic error = 16 minutes per week overestimation
Standard deviation = 806 minutes (14.3 hours)
Proportional bias = 0.11
Absolute measurement error = 562 minutes (9.4 hours) per week

Systematic error = 85 minutes per week overestimation
Standard deviation = 152 minutes
Proportional bias = 0.54
Absolute measurement error = 93 minutes per week

Systematic error = 6 minutes per week overestimation
Standard deviation = 149 minutes
Proportional bias = 0.06
Absolute measurement error = 161 minutes per week

Systematic error = 1551 minutes (25.8 hours) per week underestimation
Standard deviation = 1815 minutes (30.2 hours)
Proportional bias = 0.32
Absolute measurement error = 2002 minutes (33.4 hours) per week

Systematic error = 126 minutes per week underestimation
Standard deviation = 823 minutes (13.7 hours)
Proportional bias = 0.10
Absolute measurement error = 552 minutes (9.2 hours) per week

IPAQ: International Physical Activity Questionnaire; MVPA: moderate and vigorous activity.
References

[1] WHO. Guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization; 2020.

[2] Mercuri E, Mayhew A, Muntoni F, TREAT-NMD Neuromuscular Network, et al. Towards harmonisation of outcome measures for DMD and SMA within TREAT-NMD; report of three expert workshops: TREAT-NMD/ENMC workshop on outcome measures, 12th–13th May 2007, Naarden, The Netherlands; TREAT-NMD workshop on outcome measures in experimental trials for DMD, 30th June–1st July 2007, Naarden, The Netherlands; conjoint Institute of Myology TREAT-NMD meeting on physical activity monitoring in neuromuscular disorders, 11th July 2007, Paris, France. Neuromuscul Disord. 2008;18(11):894–903.

[3] Voet NB, van der Kooi EL, Riphagen II, et al. Strength training and aerobic exercise training for muscle disease. Cochrane Database Syst Rev. 2019;12:CD003907.

[4] Reffin J, Dewar L, and the Adult Neuromuscular Physiotherapy Special Interest Group. Exercise advice for adults with muscle-wasting conditions. In: Campagn MD, editor. www.musculardystrophyuk.org. London: Muscular Dystrophy UK; 2014.

[5] Brady L, MacNeil LG, Tarnopolsky MA. Impact of habitual exercise on the strength of individuals with myotonic dystrophy type 1. Am J Phys Med Rehabil. 2014;93(9):739–746. Quiz 747–748.

[6] Voet N, Bleijenberg G, Hendriks J, et al. Both aerobic exercise and cognitive-behavioral therapy reduce chronic fatigue in FSHD: an RCT. Neurology. 2014;83(21):1914–1922.

[7] Bankole LC, Millet GY, Temesi J, et al. Safety and efficacy of a 6-month home-based exercise program in patients with facioscapulohumeral muscular dystrophy: a randomized controlled trial. Medicine (Baltimore). 2016;95(31):e4497.

[8] Andersen G, Heje K, Buch AE, et al. High-intensity interval training in facioscapulohumeral muscular dystrophy type 1: a randomized clinical trial. J Neurol. 2017;264(6):1099–1106.

[9] Siciliano G, Simoncini C, Giannotti S, et al. Muscle exercise in limb girdle muscular dystrophies: pitfall and advantages. Acta Myol. 2015;34(1):3–8.

[10] O’Connor L, Westerberg E, Punga AR. Myasthenia gravis and physical exercise: a novel paradigm. Front Neurol. 2020;11:675.

[11] White CM, Hadden RD, Robert-Lewis SF, et al. Observer blind randomised controlled trial of a tailored home exercise programme versus usual care in people with stable inflammatory mediated neuropathy. BMC Neurol. 2015;15:147.

[12] Jones K, Hawke F, Newman J, et al. Interventions for promoting physical activity in people with neuromuscular disease. Cochrane Database Syst Rev. 2021;5:CD013544.

[13] Latouche C, Jowett JB, Carey AL, et al. Effects of breaking up prolonged sitting on skeletal muscle gene expression. J Appl Physiol (1985). 2013;114(4):453–460.

[14] Mercuri E, Mazzone E. Choosing the right clinical outcome measure: from the patient to the statistician and back. Neuromuscul Disord. 2011;21(1):16–19.

[15] Lovering RM, Porter NC, Bloch RJ. The muscular dystrophies: from genes to therapies. Phys Ther. 2005;85(12):1372–1388.

[16] Gang Q, Bettencourt C, Machado P, et al. Sporadic inclusion body myositis: the genetic contributions to the pathogenesis. Orphanet J Rare Dis. 2014;9:88.

[17] Emery AE. The muscular dystrophies. Lancet. 2002;359(9307):687–695.

[18] Mastaglia FL. Sporadic inclusion body myositis: variability in prevalence and phenotype and influence of the MHC. Acta Myol. 2009;28(2):66–71.

[19] Cassell J, Cassell R, Down F, et al. Wheelchair provision for children and adults with muscular dystrophy and other neuromuscular conditions. Best Practice Guidelines, Muscular Dystrophy Campaign. 2011:53–48.

[20] Jimenez-Moreno AC, Newman J, Charman SJ, et al. Measuring habitual physical activity in neuromuscular disorders: a systematic review. J Neuromuscul Disord. 2017;4(1):25–52.

[21] Roberts-Lewis SF, Rose MR, White CM, et al. Measurement of physical activity in adults with muscular dystrophy: a systematic review. J Phys Activity Res. 2019;4(1):1–25.

[22] Bauman A, Ainsworth BE, Bull F, et al. Progress and pitfalls in the use of the International Physical Activity Questionnaire (IPAQ) for adult physical activity surveillance. J Phys Act Health. 2009;6 Suppl 1:S55–S58.

[23] Washburn RA, Zhu W, McAuley E, et al. The physical activity scale for individuals with physical disabilities: development and evaluation. Arch Phys Med Rehabil. 2002;83(2):193–200.

[24] Hallal PC, Pratt M. Physical activity: moving from words to action. Lancet Glob Health. 2020;8(7):e867–e868.

[25] Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc. 2003;35(8):1381–1395.

[26] Hurtig-Wennlöf A, Hagströmer M, Olsson LA. The International Physical Activity Questionnaire modified for the elderly: aspects of validity and feasibility. Public Health Nutr. 2010;13(11):1847–1854.

[27] Rutkowski I, editor. Mobility in the Project – "Adapted Physical Activity – Theory and Practice". 4th International Conference: Adapted Physical Activity in Rehabilitation and Education; 2014. Warsaw: Josef Pilsudski University of Physical Education.

[28] Hagströmer M, Oja P, Sjostrom M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nutr. 2006;9(6):755–762.

[29] Ekelund U, Sepp H, Brage S, et al. Criterion-related validity of the last 7-day, short form of the International Physical Activity Questionnaire in Swedish adults. Public Health Nutr. 2006;9(2):258–265.

[30] Prince SA, Adamo KB, Hamel ME, et al. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. Int J Behav Nutr Phys Act. 2008;5:56.
