Journal of Intellectual Disability Research

ORIGINAL ARTICLE OPEN ACCESS

Feasibility, Test–Retest Reliability and Convergent Validity of the Two-Minute Step Test in Older Adults With Intellectual Disabilities

M. M. A. de Bondt[1] | T. I. M. Hilgenkamp[2,3] | A. Oppewal[3,4]

[1]Physical Therapy Sciences, Program in Clinical Health Sciences, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands | [2]Department of Physical Therapy, University of Nevada, Las Vegas, Las Vegas, Nevada, USA | [3]Academic Collaborative Research Center Healthy Ageing and Intellectual Disabilities, Rotterdam, the Netherlands | [4]Department of General Practice, Intellectual Disability Medicine Research, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands

Correspondence: A. Oppewal (a.oppewal@erasmusmc.nl)

Received: 28 April 2025 | **Revised:** 12 November 2025 | **Accepted:** 27 November 2025

Keywords: aerobic capacity | cardiorespiratory endurance | clinimetrics | instruments | intellectual development disorders | physical capability

ABSTRACT

Background: Cardiorespiratory fitness is a crucial component for health. However, measuring cardiorespiratory fitness in older adults with intellectual disabilities (ID) in practice and in large-scale studies remains challenging. The Two-Minute Step Test (2MST) is a submaximal cardiorespiratory fitness test that is brief, relatively easy to perform, and does not require any expensive materials. However, the clinimetric properties of the 2MST for older adults with ID are unknown. Therefore, this study aimed to determine the feasibility, reliability and convergent validity of the 2MST in older adults with ID.

Method: This was a cross-sectional study within the 'Healthy Ageing and Intellectual Disabilities' (HA-ID) cohort study. Participants ($n = 180$, 70.6 [66–75] years) that participated in the physical fitness assessment were included in this study. Feasibility was defined as a successful completion according to protocol. Test–retest reliability was assessed over two measurements, on the same day, with the intraclass correlation coefficient (ICC). Convergent validity was assessed relative to other fitness components (static balance, muscular endurance, comfortable and fast gait speed), with Pearson's and Spearman's correlation coefficients.

Results: The 2MST showed moderate feasibility for the first (49.4%) and second (48.3%) measurement. Feasibility was good in participants with borderline, mild and moderate ID (57.1%–80%), but low in participants with severe (15.2%) and profound ID (7.7%). Test–retest reliability was excellent (ICC = 0.88). The 2MST had a moderate correlation with static balance ($r = 0.46$) and comfortable ($r = 0.42$) and fast ($r = 0.51$) gait speed, and a good correlation with muscular endurance ($r = 0.63$). Based on predefined criteria, all correlations were at least moderate ($r > 0.30$), supporting good convergent validity (categorised as good if three or more out of four correlations were at least moderate).

Conclusions: The 2MST is a feasible, reliable and valid test to use in older adults with ID. Feasibility was especially good in older adults with borderline to moderate ID; however, it was low in adults with more severe ID. The excellent test–retest reliability and good indications for convergent validity show this test is a suitable field test for cardiorespiratory fitness to use in older adults with ID.

1 | Introduction

The general population is ageing, as are people with intellectual disabilities (ID) (Schoufour et al. 2015). Even though the life expectancy of people with ID is increasing, they often have multiple health problems at a younger age than the general population (Dolan et al. 2019; Oppewal et al. 2020; Schoufour et al. 2013). Compared with their peers in the general population, older adults with ID have multiple disabilities and chronic diseases at a younger age and become frail earlier and more severely (Reppermund and Trollor 2016; Schoufour et al. 2015). Also, older adults with ID have very poor physical fitness levels (Hilgenkamp et al. 2012a; Oppewal et al. 2020; Organization 2024) resulting in a higher risk for cardiovascular diseases, musculoskeletal conditions, decline in mobility and daily functioning and a higher mortality risk (Bull et al. 2020; de Leeuw et al. 2022; Hilgenkamp et al. 2012b; Oppewal et al. 2013; Piercy et al. 2018).

A crucial component of physical fitness is cardiorespiratory fitness. Cardiorespiratory fitness is a strong predictor for cardiovascular health problems and mortality (Abellan Van Kan et al. 2009). On the other hand, improvements in cardiorespiratory fitness are associated with improved health outcomes such as a lower risk for cardiovascular disease events (Franklin et al. 2023). Literature in the general population shows that every 1 Metabolic Equivalent of a Task (MET) increase in cardiorespiratory fitness lowers the risk for all-cause mortality by 11%–17% (Lang et al. 2024). Older adults with ID have cardiorespiratory fitness levels well below the average ranges of the general population putting them at risk for health problems and early mortality (Oppewal and Hilgenkamp 2019). Measuring cardiorespiratory fitness in older adults with ID is therefore important to identify those at risk and provide and evaluate the right care and interventions.

The gold standard test for cardiorespiratory fitness is measuring maximal oxygen uptake (VO_2 max) during a maximal exercise test (American College of Sports Medicine 2021). However, this is not always easy to perform for older adults with ID because it can be too challenging for them to complete, their physical and cognitive limitations can hinder test execution, the concept of maximal performance is not always understood and maintaining interest and motivation can be challenging (Hilgenkamp et al. 2010; Hilgenkamp et al. 2013). Additionally, a maximal exercise test is not always easy to perform in clinical practice or within large-scale studies because it requires specific expertise, sufficient familiarisation sessions and expensive materials, which are not easily transportable.

Consequently, there has been a significant effort to find suitable field tests to measure cardiorespiratory fitness in people with ID (Hilgenkamp et al. 2010; Oppewal et al. 2013; Wouters et al. 2017). The Two-Minute Step Test (2MST) is a submaximal cardiorespiratory fitness test that is brief, relatively easy to perform and requires no expensive materials. In the general population, the 2MST has a strong correlation with VO_2max (Bohannon and Crouch 2019), and is considered a valid and reliable (ICC = 0.88) test to quantify the cardiorespiratory fitness in older adults (Berlanga et al. 2023; Chen et al. 2009; Rikli and Jones 2013a). Normative reference values are available for adults aged 60–94 years (Rikli and Jones 1999). The 2MST is potentially a suitable test for older adults with ID as well, but the feasibility, reliability and validity of the 2MST for this group are unknown. Therefore, this study aims to determine the feasibility, reliability and convergent validity of the 2MST in older adults with ID.

2 | Methods

2.1 | Study Design and Participants

This was a cross-sectional study within the 'Healthy Ageing and Intellectual Disabilities' (HA-ID) prospective multicentre cohort study (De Leeuw et al. 2022), performed by the Academic Collaborative Research Center HA-ID. This is a partnership between three ID healthcare organisations (Abrona, Amarant and Ipse de Bruggen) in the Netherlands and the Department of General Practice, Intellectual Disability Medicine Research group of the Erasmus MC, University Medical Center Rotterdam. In 2008, all older adults with ID aged 50 years or older, receiving care or support from one of the HA-ID care organisations ($n = 2322$) were invited to participate, resulting in a sample of 1050 participants (Hilgenkamp et al. 2011). All 429 participants who still received care from one of the HA-ID care organisations on 1 July 2020 were invited to participate in the 10-year follow-up measurements, of whom 278 participants provided consent (De Leeuw et al. 2022). If participants were incapable of providing informed consent themselves, informed consent was provided by their legal representative. Data were collected between October 2020 and July 2023.

To be eligible for this study participants had to meet the following criteria: participating in the 10-year follow-up physical fitness assessment of the HA-ID study and at least attempt to perform the 2MST. At the time of the 10-year follow-up all participants were 60 years or older. To assure safety during testing, the Physical Activity Readiness Questionnaire (PAR-Q) was filled out prior to testing by professional caregivers (Thomas et al. 1992). If any of the questions were answered with 'yes' or 'unknown', the physician of the participant was consulted to determine whether it was safe for the participant to perform the 2MST.

Ethical approval was obtained from the Medical Ethics Review Committee of the Erasmus MC, University Medical Center Rotterdam (MEC-2019-0562). The study is registered in the 'Overview of Medical Research in the Netherlands' (OMON; NL-OMON28611, https://onderzoekmetmensen.nl/nl/trial/28611), which is linked to the International Clinical Trial Registry Platform (ICTRP). The study follows the guidelines of the Declaration of Helsinki (World Medical Association 2025).

2.2 | Measurements

2.2.1 | Personal Characteristics

Baseline data on age and sex were collected from the administrative systems. Level of ID was collected from psychologists' and behavioural therapists' files, categorised as borderline

(IQ = 70–84), mild (IQ = 50–69), moderate (IQ = 35–59), severe (IQ = 20–34) and profound (IQ < 20). Participants' residential status (central setting, community-based, independent with ambulatory support) and mobility (independent, walking with an aid, wheelchair-bound) were collected through questionnaires filled out by the professional caregiver of the participant. Body mass index (BMI) was calculated based on height and weight measured during physical examination.

2.2.2 | Cardiorespiratory Fitness

Cardiorespiratory fitness was measured with the 2MST (Bohannon and Crouch 2019; Rikli and Jones 2001). The 2MST is considered a valid and reliable (ICC = 0.88) test to measure cardiorespiratory fitness in older adults in the general population (Berlanga et al. 2023; Chen et al. 2009; Rikli and Jones 2013b). Participants had to raise their knees as often as possible to a predetermined height within 2 min. First, step height was determined per participant at the centre between the iliac crest and the patella. This height was marked on a wall. The number of times the correct step height was reached with the right knee was counted, and this was the score on the test. If the participant was not able to reach the required step height, for example because of limited range of motion, adjustments were made to the height. The adjustment was made to the maximum height that the participant could reach at that moment (Rikli and Jones 2001). If with these adjustments the test could be performed successfully, this was considered a feasible performance. Participants were encouraged to provide maximal effort. To assure safety, we monitored for signs of overexertion. Reasons for a non-successful performance were noted (concentration problems, difficulties understanding the task, physical limitations, lack of motivation, anxiety, other, unknown).

2.2.3 | Other Physical Fitness Components

For convergent validity, we selected those fitness tests from the HA-ID physical fitness assessment most closely related to cardiorespiratory fitness.

Static balance was measured with the balance stances from the Short Physical Performance Battery (SPPB) (Guralnik et al. 1994). Participants had to complete three stances with increasing difficulty: side-by-side, semi-tandem and full-tandem stand. Each stance had to be held for 10 s. The scoring system of the SPPB was used, with scores ranging from 0 (not able to perform any of the stances for 10 s) to 4 (able to perform all stances for 10 s) (Guralnik et al. 1994). The test–retest reliability of the SPPB within community-dwelling older adults is good (Olsen and Bergland 2017).

Muscular endurance was measured with the 30s Chair stand (Jones et al. 1999). Participants had to stand up from a chair and sit down again as often as possible within 30 s without using their arms (Jones et al. 1999). The number of complete stances was the test result. In older adults with ID, feasibility and test–retest reliability were moderate (Hilgenkamp et al. 2012b; Hilgenkamp et al. 2013).

Comfortable and maximal gait speed was measured over a 5-m distance, after 3 m for acceleration and ending with 3 m for deceleration (Bohannon 1997). Participants walked three times for each condition. For both gait speeds, the average of the three trials was the test result (m/s). Participants were allowed to use their walking aid when needed. In older adults with ID, feasibility and test–retest reliability were good (Berlanga et al. 2023; Chen et al. 2009; Hilgenkamp et al. 2013; Rikli and Jones 1999).

2.2.4 | Procedures

The fitness tests included in this study were part of a more comprehensive physical fitness assessment, lasting for about 90 min total. The first 2MST was performed relatively at the start and the second one at the end of the fitness assessment. Participants were given a minimum of 5 min rest to ensure they were adequately rested before the second 2MST began.

The physical fitness assessment took place at locations familiar or close to the participant. The tests were administered in a predetermined order and conducted by trained physiotherapists, physical activity instructors or occupational therapists with several years of experience in working with people with ID. The test instructor was instructed to maximally motivate the participant and to only record the test result as valid if they were convinced that participants understood the task and performed the 2MST with maximal effort and according to protocol. The experience of the test administrators was important to ensure the most suitable 'maximal motivation' for each participant, regarding safety as well. Although no formal inter-rater reliability assessment was performed, the test instructors' prior experience and training ensured consistency in application and judgement for a successful performance.

2.3 | Statistical Analyses

Descriptive statistics of personal characteristics were provided for the total study sample, and the sample that had sufficient data for the reliability and validity analyses. Depending on the normality of the data, results were reported as means with standard deviations or medians with interquartile ranges.

2.3.1 | Feasibility

The feasibility of the 2MST was quantified as the percentage of participants able to perform the 2MST according to protocol. The 2MST execution was considered unsuccessful when the participant was not able to meet the required step height and did not finish the test because of other reasons than being too exhausted to finish the 2 min. Feasibility was assessed for the total group, as well as for subgroups based on age, sex, level of ID and mobility. Results are presented for the first and second measurements. Feasibility was classified as low (≤ 25%), moderate (> 25% and ≤ 50%), good (> 50% and ≤ 75%) and excellent (> 75%) (Hilgenkamp et al. 2013). Reasons for non-successful performance were described.

2.3.2 | Test–Retest Reliability

To be included in the reliability analysis, participants had to complete both 2MST measurements. Test–retest reliability was analysed with the Interclass Correlation Coefficient (ICC), using a two-way mixed model. The ICC was categorised as low (0 and ≤0.300), moderate (between >0.300 and ≤0.500), good (>0.500 and ≤0.750) and excellent (>0.750) (Cohen 1988; Hilgenkamp et al. 2013).

2.3.3 | Convergent Validity

To be included in the validity analyses, participants had to have completed the other physical fitness tests as well as the 2MST. First, we analysed the amount of missing data. Within the complete case analysis there was a maximum percentage of 4.8% missing data. Research has shown that missing data below 5% gives a low chance of bias (Michy 2015; Schafer 1999), and therefore we decided not to apply missing data procedures. The convergent validity of the 2MST was assessed in relation to static balance, muscular endurance and gait speed. Depending on the normality of the data, either Pearson's or Spearman's correlation was used. The first measurement of the 2MST was used for the convergent validity. This was done because of the anticipated presence of a potential learning effect within the 2MST. All correlations were characterised as low (0 and ≤0.30), moderate (>0.30 and ≤0.50), good (>0.50 and ≤0.75) and excellent (>0.75) (Hilgenkamp et al. 2013). The convergent validity was categorised as moderate if two out of four correlations were at least moderate, and as good if three or more out of four correlations were at least moderate.

All analyses were performed with the Statistical Package for Social Sciences 28.0 (SPSS 28.0) (IBM Corporation, New York). Statistical significance was set at 5% ($p<0.05$).

3 | Results

3.1 | Personal Characteristics

Table 1 presents the personal characteristics for the total study sample ($n=180$), and for the group included in the reliability and convergent validity analyses ($n=84$). For the total study sample, the median age was 70.61 (66–75) years, 49.4% was female and 38.9% was independent in mobility. More than half of the sample was living in a central care facility (52.8%). The subgroup that was included in the reliability and convergent validity analyses had less severe ID, and more often lived in a community-based setting.

3.2 | Feasibility

The feasibility results of both 2MST measurements are presented in Table 2. For both measurements, feasibility was moderate (49.4% and 48.3%). Feasibility was lowest in older adults with severe (15.2%) and profound ID (7.7%). Older adults in a wheelchair were unable to perform the 2MST and therefore feasibility was low in this subgroup. Regarding mobility, feasibility was good in both older adults who used a walking aid and who were independent in mobility, except in the second 2MST where feasibility was moderate in adults who were independent in mobility. Reasons for a non-successful performance were difficulties understanding the task (first 2MST 41.8%, second 2MST 37.6%), physical limitations (first 2MST 31.9%, second 2MST 33.3%), lack of motivation (first 2MST 5.5%, second 2MST 7.5%), did not exert him/herself (both 2MST's 3.3%), concentration problems (both 2MST's 2.2%), or anxiety (both 2MST's 1.1%).

3.3 | Test–Retest Reliability

The ICC for test–retest reliability of the 2MST was 0.884 ($p<0.001$) with a 95% confidence interval of 0.672–0.946, indicating excellent test–retest reliability.

3.4 | Convergent Validity

The correlations between the 2MST and other physical fitness tests are shown in Table 3. A higher 2MST score significantly correlated with higher static balance (r=0.46; moderate), higher muscular endurance (r=0.63; good), higher comfortable gait speed (r=0.42; moderate) and maximal gait speed (r=0.51; good). With all correlations being at least moderate, these results demonstrate good convergent validity.

4 | Discussion

This study is the first to explore the feasibility, reliability and convergent validity of the 2MST in older adults with ID. The overall feasibility of the 2MST was moderate (49.4 and 48.3%). Feasibility was good in participants with borderline, mild and moderate ID (57.1%–80%), but low in participants with severe and profound ID (7.7%–15.2%). The test–rest reliability of the 2MST was excellent (ICC=0.886), and convergent validity was good as supported by all four fitness tests presenting a moderate to good correlation with the 2MST.

The overall moderate feasibility suggests that the 2MST can successfully be used within older adults with ID. However, the level of ID impacted feasibility, with the lowest feasibility for those with severe (15.2%) and profound ID (7.7%). This is likely associated with more difficulties understanding the task and more physical limitations in participants with more severe ID, which were the main reasons for non-successful performance.

The 2MST showed excellent reliability, indicating consistency in test–retest measurements (Michalos 2014). The test–retest reliability in this study (ICC=0.884) was similar to that found in healthy older adults (ICC=0.880) (Rikli and Jones 1999; Rikli and Jones 2013a). Although reliability was excellent, we did see a better score on the second measurement, indicating a learning effect. This learning effect has also been found previously (Chow et al. 2023), therefore it is recommended to practice the test with a participant before the actual testing.

The 2MST also showed good convergent validity, with good correlations with muscular endurance (r=0.63) and maximal gait

TABLE 1 | Personal characteristics and results on the Two-Minute Step Test presented for the total study population, and the subgroup included in the reliability and validity analyses.

	Total sample $n = 180$	Subgroup reliability & validity analyses $n = 84$
Age (median, IQR)	70.61 (66–75)	70.58 (66–75)
Age categories (*n*, %)		
60–65 years	42 (23.3%)	18 (21.4%)
66–70 years	61 (33.9%)	30 (35.7%)
71–75 years	35 (19.4%)	17 (20.2%)
76–80 years	32 (17.8%)	14 (16.7%)
80+ years	10 (5.6%)	5 (6.0%)
Sex (*n*, %)		
Female	89 (49.4%)	42 (50.0%)
BMI (mean ± SD)	26.40 ± 5.15	27.16 ± 5.16
Level of ID (*n*, %)		
Borderline	10 (5.6%)	7 (8.3%)
Mild	32 (17.8%)	21 (25.0%)
Moderate	91 (50.6%)	51 (60.7%)
Severe	33 (18.3%)	4 (4.8%)
Profound	13 (7.2%)	1 (1.2%)
Unknown	1 (0.6%)	0 (0.0%)
Mobility (*n*, %)		
Independent	70 (38.9%)	33 (39.3%)
Walking-aid	34 (18.9%)	23 (27.4%)
Wheelchair	15 (8.3%)	0 (0.0%)
Unknown	21 (11.7%)	8 (9.5%)
Missing value	40 (22.2%)	20 (23.8%)
Residential status (*n*, %)		
Central setting	95 (52.8%)	28 (33.3%)
Community based	81 (45.0%)	52 (61.9%)
Independent with ambulatory support	1 (0.6%)	1 (1.2%)
Unknown	3 (1.7%)	3 (3.6%)
2MST-1 (median, IQR)	41.12 (23–55) $n = 89$	41.13 (23–56) $n = 84$
2MST-2 (mean ± SD)	46.49 ± 24.65 $n = 87$	47.60 ± 24.28 $n = 84$

Note: %: Percentages indicate the proportion of the study population that falls within each specific category or subgroup.
Abbreviations: 2MST-1: first Two-Minute Step Test; 2MST-2: second Two-Minute Step Test; IQR: inter quartile range; *n*: number of participants, SD: standard deviation.

speed (r = 0.51), and moderate correlations with static balance (r = 0.46) and comfortable gait speed (r = 0.42). Comparing our results with other studies is difficult because not a lot of studies have calculated these correlations or use different tests for the fitness components. For muscular endurance, a correlation of r = −0.405 between the 5-times Chair stand and the 2MST in adults with Multiple Sclerosis was found and a correlation of r = −0.871 in adults with Diabetes mellitus type 2 (Özkeskin et al. 2023; Srithawong et al. 2022). The correlation of r = 0.63 between muscular endurance (30s Chair stand) and the 2MST found in our study falls within these two correlations. For the 30s Chair stand test, a better performance is reflected by a

TABLE 2 | Feasibility results for the first and second 2MST measurements.

	2MST-1 (% of row) ($n=180$)		2MST-2 (% of row) ($n=180$)	
Total sample	49.4	Mod.	48.3	Mod.
Age				
60–65 years ($n=42$)	45.2	Mod.	42.9	Mod.
66–70 years ($n=61$)	55.7	Good	50.8	Mod.
71–75 years ($n=35$)	48.6	Mod.	48.6	Mod.
76–80 years ($n=32$)	43.8	Mod.	46.9	Mod.
80+ years ($n=10$)	50.0	Good	60.0	Good
Sex				
Female ($n=89$)	49.4	Mod.	49.4	Mod.
Male ($n=91$)	49.5	Mod.	47.3	Mod.
Level of ID				
Borderline ($n=10$)	80.0	Exc.	70.0	Good
Mild ($n=32$)	68.8	Good	68.8	Good
Moderate ($n=91$)	58.2	Mod.	57.1	Good
Severe ($n=33$)	15.2	Low	15.2	Low
Profound ($n=13$)	7.7	Low	7.7	Low
Mobility				
Independent ($n=70$)	51.4	Good	47.1	Mod.
Walking-aid ($n=34$)	67.6	Good	73.5	Good
Wheelchair ($n=15$)	0	Low	0	Low

Note: Feasibility categories: low (≤25%), moderate (>25% and ≤50%), good (>50% and ≤75%) and excellent feasibility (>75%).
Abbreviations: 2MST-1: first Two-Minute Step Test; 2MST-2: second Two-Minute Step Test; Exc.: excellent; Mod.: moderate.

TABLE 3 | Correlations between the 2MST and other physical fitness tests.

Physical fitness component	*n*	Correlation	95% CI Lower	95% CI Upper
Static balance	84	0.46[b]	0.27	0.62
Muscular endurance	81	0.63[a]	0.48	0.75
Comfortable gait speed[c]	84	0.42[b]	−0.58	−0.22
Maximal gait speed[c]	80	0.51[b]	−0.66	−0.32

Abbreviations: CI: confidence interval; *n*: number of participants.
[a]Pearson correlation.
[b]Spearman correlation.
[c]A lower score representing a better performance.

greater number of repetitions and for the 5-times Chair stand test, a better performance is reflected by a shorter completion time, explaining the negative and positive correlations.

This study had some strengths and limitations. One of the strengths is the relatively large sample size of older adults with ID, and the advanced age of the included participants. Another strength is the large number of physical fitness tests performed, providing objective data and allowing for comparison of the 2MST with a wide range of physical fitness components. However, some selection bias possibly exists because of the complete case analysis conducted for the reliability and validity analyses. This subgroup included for these analyses consisted of participants with less severe ID, who more often lived in supported community settings than in central care facilities. Similarly, our study sample ($n=180$) was a subset of the participants ($n=278$) that provided consent for this wave of the HA-ID cohort study (De Leeuw et al. 2022). Some participants only provided consent for file research using the existing health records, and not for active participation in the physical measurements. Additionally, not everybody was able to participate in the physical fitness measurements. Therefore, the study sample of this study might represent the better functioning group of the total HA-ID study sample. Also, the HA-ID study population does not include older adults with ID who do not receive any form of registered formal support or care and are living independently. Therefore, the results are not generalisable to this group. Considering the trend evident in the presented results, which indicates a more positive outlook for subgroups characterised by higher functional and cognitive ability, it is conceivable that these same instruments can be applied to measure cardiorespiratory fitness in adults with ID living independently and of younger age.

Further research is needed to explore the potential learning effect of the 2MST. This study showed that on average the score on the second measurement was better than on the first. Because of this potential learning effect, it is important to investigate whether the initial administration of the 2MST accurately assesses cardiorespiratory fitness. Additionally, future studies should examine the validity of the 2MST against the gold standard, VO_2 max obtained from a maximal exercise test, in adults with ID and longitudinal research is needed to study the predictive validity of the 2MST for health issues. Finally, for individuals with severe or profound ID, adaptations to the testing procedure might improve feasibility, such as (video) modelling of movement. Further research is needed to make the 2MST more suitable for adults with more severe ID, or to develop and validate other suitable testing methods for this subgroup.

5 | Conclusion

This study shows that the 2MST is a test that can be used to assess cardiorespiratory fitness in older adults with ID, with moderate feasibility, excellent reliability and good convergent validity. Based on this study, we recommend the 2MST for the group with borderline to moderate ID who are able to walk independently or with walking aids, while feasibility was low for those with severe and profound ID and wheelchair users.

In conclusion, this study fills a crucial gap in clinical practice regarding cardiorespiratory fitness testing, by showing moderate feasibility, excellent reliability and good convergent validity for the 2MST.

Acknowledgements

The authors thank the care organisations involved in the Academic Collaborative Research Center Healthy Ageing and Intellectual Disabilities (HA-ID), Abrona, Amarant and Ipse de Bruggen, for their collaboration and financial and organisational support. We also thank the professionals of these organisations as well as all participants, their family members and caregivers for their invaluable contribution to the HA-ID study. Lastly the authors thank the lecturers of the study Clinical Health Science of the University of Utrecht for their feedback on this article.

Funding

The Academic Collaborative Research Center Healthy Ageing and Intellectual Disabilities (HA-ID) funded the HA-ID cohort study. This consortium consists of three Dutch care organisations for people with intellectual disabilities (Abrona, Amarant and Ipse de Bruggen) in collaboration with the Department of General Practice, Intellectual Disability Medicine Research of the Erasmus MC, University Medical Center Rotterdam. The three care organisations were not involved in the data processing, analysis and writing of the manuscript.

Ethics Statement

Ethical approval was obtained from the Medical Ethics Review Committee of the Erasmus MC, University Medical Center Rotterdam (MEC-2019-0562). The study is registered in the Dutch Trial Register (NTR number: NL8564, https://onderzoekmetmensen.nl/nl/trial/28611) and follows the guidelines of the Declaration of Helsinki (World Medical Association 2025).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Abellan Van Kan, G., Y. Rolland, S. Andrieu, et al. 2009. "Gait Speed at Usual Pace as a Predictor of Adverse Outcomes in Community-Dwelling Older People an International Academy on Nutrition and Aging (IANA) Task Force." *Journal of Nutrition, Health & Aging* 13: 881–889. https://doi.org/10.1007/s12603-009-0246-z.

American College of Sports Medicine. 2021. *ACSM'S Guidelines for Exercise Testing and Prescription*. 11th ed. Wolters Kluwer.

Berlanga, L. A., M. Matos-Duarte, P. Abdalla, E. Alves, J. Mota, and L. Bohn. 2023. "Validity of the Two-Minute Step Test for Healthy Older Adults." *Geriatric Nursing* 51: 415–421.

Bohannon, R. W. 1997. "Comfortable and Maximum Walking Speed of Adults Aged 20-79 Years: Reference Values and Determinants." *Age and Ageing* 26: 15–19.

Bohannon, R. W., and R. H. Crouch. 2019. "Two-Minute Step Test of Exercise Capacity: Systematic Review of Procedures, Performance, and Clinimetric Properties." *Journal of Geriatric Physical Therapy* 42: 105–112.

Bull, F. C., S. S. Al-Ansari, S. Biddle, et al. 2020. "World Health Organization 2020 Guidelines on Physical Activity and Sedentary Behaviour." *British Journal of Sports Medicine* 54: 1451–1462.

Chen, H. T., C. H. Lin, and L. H. Yu. 2009. "Normative Physical Fitness Scores for Community-Dwelling Older Adults." *Journal of Nursing Research* 17: 30–41.

Chow, J. J. L., C. Fitzgerald, and S. Rand. 2023. "The 2 min Step Test: A Reliable and Valid Measure of Functional Capacity in Older Adults Post Coronary Revascularisation." *Physiotherapy Research International* 28: e1984.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum.

De Leeuw, M. J., A. Oppewal, R. G. Elbers, et al. 2022. "Healthy Ageing and Intellectual Disability Study: Summary of Findings and the Protocol for the 10-Year Follow-Up Study." *BMJ Open* 12: e053499.

Dolan, E., J. Lane, G. Hillis, and N. Delanty. 2019. "Changing Trends in Life Expectancy in Intellectual Disability Over Time." *Irish Medical Journal* 112: 1006.

Franklin, B. A., I. J. Wedig, R. E. Sallis, C. J. Lavie, and S. J. Elmer. 2023. "Physical Activity and Cardiorespiratory Fitness as Modulators of Health Outcomes: A Compelling Research-Based Case Presented to the Medical Community." *Mayo Clinic Proceedings* 98: 316–331.

Guralnik, J. M., E. M. Simonsick, L. Ferrucci, et al. 1994. "A Short Physical Performance Battery Assessing Lower Extremity Function: Association With Self-Reported Disability and Prediction of Mortality and Nursing Home Admission." *Journal of Gerontology* 49: M85–M94.

Hilgenkamp, T. I., L. P. Bastiaanse, H. Hermans, C. Penning, R. Van Wijck, and H. M. Evenhuis. 2011. "Study Healthy Ageing and Intellectual Disabilities: Recruitment and Design." *Research in Developmental Disabilities* 32: 1097–1106.

Hilgenkamp, T. I., R. Van Wijck, and H. M. Evenhuis. 2010. "Physical Fitness in Older People With ID-Concept and Measuring Instruments: A Review." *Research in Developmental Disabilities* 31: 1027–1038.

Hilgenkamp, T. I., R. Van Wijck, and H. M. Evenhuis. 2012a. "Feasibility and Reliability of Physical Fitness Tests in Older Adults With Intellectual Disability: A Pilot Study." *Journal of Intellectual & Developmental Disability* 37: 158–162.

Hilgenkamp, T. I., R. Van Wijck, and H. M. Evenhuis. 2012b. "Low Physical Fitness Levels in Older Adults With ID: Results of the HA-ID Study." *Research in Developmental Disabilities* 33: 1048–1058.

Hilgenkamp, T. I., R. Van Wijck, and H. M. Evenhuis. 2013. "Feasibility of Eight Physical Fitness Tests in 1,050 Older Adults With Intellectual Disability: Results of the Healthy Ageing With Intellectual Disabilities Study." *Intellectual and Developmental Disabilities* 51: 33–47.

Jones, C. J., R. E. Rikli, and W. C. Beam. 1999. "A 30-S Chair-Stand Test as a Measure of Lower Body Strength in Community-Residing Older Adults." *Research Quarterly for Exercise and Sport* 70: 113–119.

Lang, J. J., S. A. Prince, K. Merucci, et al. 2024. "Cardiorespiratory Fitness Is a Strong and Consistent Predictor of Morbidity and Mortality Among Adults: An Overview of Meta-Analyses Representing Over 20.9 Million Observations From 199 Unique Cohort Studies." *British Journal of Sports Medicine* 58: 556–566.

Michalos, A. C. 2014. *Encyclopedia of Quality of Life and Well-Being Research*. Springer. https://doi.org/10.1007/978-94-007-0753-5.

Michy, A. 2015. "*Imputing Missing Data With R; MICE Package* [Online]". DataScience+. Available: https://datascienceplus.com/imputing-missing-data-with-r-mice-package/ [Accessed 2024].

Olsen, C. F., and A. Bergland. 2017. "Reliability of the Norwegian Version of the Short Physical Performance Battery in Older People With and Without Dementia." *BMC Geriatrics* 17: 124.

Oppewal, A., and T. I. M. Hilgenkamp. 2019. "Physical Fitness Is Predictive for 5-Year Survival in Older Adults With Intellectual

Disabilities." *Journal of Applied Research in Intellectual Disabilities* 32: 958–966.

Oppewal, A., T. I. Hilgenkamp, R. Van Wijck, and H. M. Evenhuis. 2013. "Cardiorespiratory Fitness in Individuals With Intellectual Disabilities—A Review." *Research in Developmental Disabilities* 34: 3301–3316.

Oppewal, A., D. Maes-Festen, and T. I. M. Hilgenkamp. 2020. "Small Steps in Fitness, Major Leaps in Health for Adults With Intellectual Disabilities." *Exercise and Sport Sciences Reviews* 48: 92–97.

Organization W. H. 2024. "*Ageing and health* [Online]". Available: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [Accessed 2024].

Özkeskin, M., F. Özden, E. Ar, and N. Yüceyar. 2023. "The Reliability and Validity of the 30-Second Chair Stand Test and Modified Four Square Step Test in Persons With Multiple Sclerosis." *Physiotherapy Theory and Practice* 39: 2189–2195.

Piercy, K. L., R. P. Troiano, R. M. Ballard, et al. 2018. "The Physical Activity Guidelines for Americans." *Journal of the American Medical Association* 320: 2020–2028.

Reppermund, S., and J. N. Trollor. 2016. "Successful Ageing for People With an Intellectual Disability." *Current Opinion in Psychiatry* 29: 149–154.

Rikli, R. E., and C. J. Jones. 1999. "Functional Fitness Normative Scores for Community-Residing Older Adults, Ages 60–94." *Journal of Aging and Physical Activity* 7: 162–181.

Rikli, R. E., and C. J. Jones. 2001. *Senior Fitness Test Manual*. Human Kinetics.

Rikli, R. E., and C. J. Jones. 2013a. "Development and Validation of Criterion-Referenced Clinically Relevant Fitness Standards for Maintaining Physical Independence in Later Years." *Gerontologist* 53: 255–267.

Rikli, R. E., and C. J. Jones. 2013b. *Senior Fitness Test Manual*. Human Kinetics.

Schafer, J. L. 1999. "Multiple Imputation: A Primer." *Statistical Methods in Medical Research* 8: 3–15.

Schoufour, J. D., A. Mitnitski, K. Rockwood, H. M. Evenhuis, and M. A. Echteld. 2013. "Development of a Frailty Index for Older People With Intellectual Disabilities: Results From the HA-ID Study." *Research in Developmental Disabilities* 34: 1541–1555.

Schoufour, J., H. Evenhuis, and M. Echteld. 2015. "Kwetsbaarheid bij ouderen met een verstandelijke beperking: operationalisering, risico en opsporing." *Tijdschrift voor Gerontologie en Geriatrie* 46: 92–103. https://doi.org/10.1007/s12439-015-0126-4.

Srithawong, A., P. Poncumhak, P. Manoy, et al. 2022. "The Optimal Cutoff Score of the 2-Min Step Test and Its Association With Physical Fitness in Type 2 Diabetes Mellitus." *Journal of Exercise Rehabilitation* 18: 214–221.

Thomas, S., J. Reading, and R. J. Shephard. 1992. "Revision of the Physical Activity Readiness Questionnaire (PAR-Q)." *Canadian Journal of Sport Sciences* 17: 338–345.

World Medical Association. 2025. "World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Participants." *Journal of the American Medical Association* 333, no. 1: 71–74.

Wouters, M., H. M. Evenhuis, and T. I. Hilgenkamp. 2017. "Systematic Review of Field-Based Physical Fitness Tests for Children and Adolescents With Intellectual Disabilities." *Research in Developmental Disabilities* 61: 77–94.