Volume 42 • Number 2 • December 2022
www.worldscientific.com/worldscinet/hkpj

HONG KONG PHYSIOTHERAPY JOURNAL

World Scientific

HONG KONG PHYSIOTHERAPY JOURNAL

Volume 42 • Number 2 • December 2022

www.worldscientific.com/worldscinet/hkpj

Editor-in-Chief

Professor Marco Y.C. Pang, PhD

Department of Rehabilitation Sciences,
The Hong Kong Polytechnic University, Hong Kong

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Hong Kong Physiotherapy Journal

Aims & Scope

The Hong Kong Physiotherapy Journal (HKPJ) is the official peer-reviewed, Open Access (OA) publication of the Hong Kong Physiotherapy Association.

HKPJ publishes papers related to all areas of physiotherapy (education, research, practice, policies) and is committed to facilitating communication among educators, researchers and practitioners in the field with the aim of promoting evidence-based practice.

We are particularly interested in publishing randomized controlled trials, systematic reviews and meta-analyses. Animal studies are also welcome if the study question and findings have important relevance to physiotherapy practice.

HKPJ welcomes submissions from all over the world in the form of original research papers, reviews, editorials, treatment reports, technical notes, and correspondence.

Copyright © 2022 Hong Kong Physiotherapy Association

Hong Kong Physiotherapy Journal [ISSN (print): 1013-7025 | ISSN (online): 1876-441X] is published by World Scientific Publishing Co Pte Ltd. This is an Open Access journal. The articles are published under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

HONG KONG PHYSIOTHERAPY JOURNAL

Vol. 42, No. 2, December 2022

Technical Report

Hong Kong Physiotherapy Journal
Vol. 42, No. 2 (2022) 75–80
DOI: 10.1142/S1013702522710019

Hong Kong Physiotherapy Journal
https://www.worldscientific.com/worldscinet/hkpj

A technical report on a novel robotic lower limb rehabilitation device - Is ROBERT® a cost-effective solution for rehabilitation in Hong Kong?

Alan Chan

Faculty of Health and Social Sciences, The Hong Kong Polytechnic University
Hong Kong SAR, China

Li Ka Shing Faculty of Medicine, The University of Hong Kong
Hong Kong SAR, China

alan_chan_hk@hotmail.com.hk

Received 25 March 2022; Accepted 11 July 2022; Published 24 August 2022

A shortage of physiotherapist (PT) manpower is a barrier for providing better rehabilitation service in Hong Kong. Quality training can benefit patients with better recovery, on the contrary, insufficient training may cause a longer length of stay, readmission, and thus the burden of healthcare system. The estimated cost for PT services in Hospital Authority was HK$7.0 Billion in 2020. A novel Danish robot with a 7-joint robotic arm became popular in Denmark and Germany in the last two years. The robot is designed for lower limb patient rehabilitation. It can enhance the mobility of patients. Based on the experience of a university hospital in Denmark, this robotic rehabilitation was well accepted by both patients and PTs. Function-wise, the robot provides many clinical benefits to patients, especially stroke ones. A physiotherapist's time can be saved when the robot is being used. The cost-effectiveness of ROBERT® is better than PT performing repetitive exercises for lower limbs. The robot potentially provides a cost-effective solution to the Hong Kong healthcare system.

Keywords: Robot; ROBERT; stroke; rehabilitation; length of stay; effectiveness.

Copyright@2022, Hong Kong Physiotherapy Association. This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) License which permits use, distribution and reproduction, provided that the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Background

Physiotherapists (PTs) diagnose and assess injured patients in physical means. Their role is exceptionally essential to stroke rehabilitation. Daras *et al.* stated that with additional physical activities during Length of Stay (LOS), patients can have better functional independence during discharge and thus lower the readmission opportunity.[1] A long LOS of stroke patients results in hospital bed occupation in the Hong Kong Hospital Authority (HA) hospitals. To meet the stroke rehabilitation goals, a guideline from the Royal College of Physicians of United Kingdom recommended that exercise intensity should accumulate to at least 45 min every day.[2] Slade *et al.* conducted a single-blinded research to examine the effect of intensity of physiotherapy upon LOS for 126 stroke patients. The control group received 59 min exercise daily while the Quality Improved (QI) group received 77 min. The mean LOS for all patients was 84.6 days, but the QI group resulted in 14 days LOS (or 16.5%) shorter than the control group with statistical significance.[3] Many bedridden patients, who suffer from muscle wasting, are susceptible to increase fall risk. Improved post-stroke patient physical condition enhances mobility and reduces opportunity of complications, e.g., falls and pressure injury. On the contrary, poor physical condition would prolong LOS and increase the potential of readmission in future, which lead to burdening of the healthcare system.

Concern: Burden in Hong Kong Physiotherapy

To the best of our knowledge about acute stroke case number in Hong Kong, there were about 16,900 cases annually between year 1999 and 2007 including new and recurrent cases, which was from Prof. J. Woo's research team articles.[4,5] The mean LOS of 1,111 stroke patients in Kowloon Hospital was 36.6 days.[6] The HA Annual Report 2020–2021 showed that the cost for acute or convalescent inpatient service was HK$ 7,240 per patient day.[7] With a direct calculation with the above figures, the estimated cost paid by HA for stroke inpatients was HK$ 4.48 billion in 2020.

Apart from post-stroke rehabilitation, PTs also treat other physically injured patients. They served about 1.2 million and 3.0 million PT patient attendances out- and in-patients, respectively, in HA hospitals in 2020.[8] The average session cost for specialist outpatient attendance among HA hospitals was HK$1,660.[7] Assuming the cost of the PT inpatient service is the same as that of outpatient, the estimated cost for PT service in 2020 in HA was close to HK$7.0 billion. The expenditure on Physiotherapy treatments is considerable. An effective and efficient rehabilitation helps HA save costs and reduce the economic burden.

Cause: PT Manpower Shortage

According to the HA Annual Report 2020–2021, there were 1,248 PTs in HA.[7] With a calculation of an ideal case that all 1,248 PTs performing exercise at full throttle, 9 h a day, Monday to Saturday, to 4.2 million in-/out-patient attendances. On average, every attendance receives around 0.72 h or 43.2 min per session by PT. In reality, PTs have to spend time on administration, treatment record documentation, research works, etc. No PT can perform exercise non-stopped for 9 h a day. Some exercises are performed by assistants inevitably. Are the treatments similar in quality? Are PT session durations sufficient in quantity?

Shortage of PT is the major barrier leading PT not being able to provide quality exercise to patients. The Standard dated on 19 June 2017 reported that "*The projection done for the government indicated that Hong Kong could see a shortage of 933 physiotherapists by 2030. But two unions and a pressure group from the physiotherapy sector said yesterday the figure is a serious underestimation.*[9]" Currently, there are 3,950 registered practicing PTs on the registration list of the Hong Kong Physiotherapists Board. Among them, about 2,700 PTs are in private practice whereas 1,250 PTs are in HA. Considering the aging population in Hong Kong, physiotherapy is facing a severe manpower shortage issue.

Solution: Robot Therapy

In 2013, Kwok and Ma foresaw that rehabilitation training would be robot-driven in future. Robots can provide more exercises to patients with minimum PT manpower. It saves PT time and assists PTs gather treatment parameters to prepare a tailor-made training.[10] Lum *et al.* and Hug *et al.* recognised that robot provides a cost-effective means for stroke patients to maintain mobility.[11,12] Few criteria were mentioned by Lum *et al.* to

choose a successful Robot therapy system, i.e., the device

(1) must provide quantifiable, functional benefits to patients,
(2) should improve the efficiency of therapists' current practice,
(3) should be affordable,
(4) should not increase the cost of health care.[11]

A novel robotic device — ROBERT®

Bertelsen *et al.* shared their experience of a pilot test in geriatric wards of Odense University Hospital with a Danish robot called ROBERT® (Life Science Robotics, Aalborg, Denmark). The research employed the passive mobilization of the robot to treat 13 elderly patients. Bertelsen *et al.* concluded that the pilot test was well accepted by patients, relatives, and staff.[13] The system overview is as shown in Fig. 1.

ROBERT® Technical Capabilities

The Danish robot is designed for performing lower limb exercises with a 7-joint robotic arm technology which allows a wide range of movement. The robot can lift an injured leg of maximum 11 kg which is equivalent to a patient body mass of approximately of 160 kg. The PT can move the impaired leg with little effort because the weight of the leg is compensated. With the excellent movement feasibility of the robotic arm, one PT alone can perform Proprioceptive Neuromuscular Facilitation training easily.

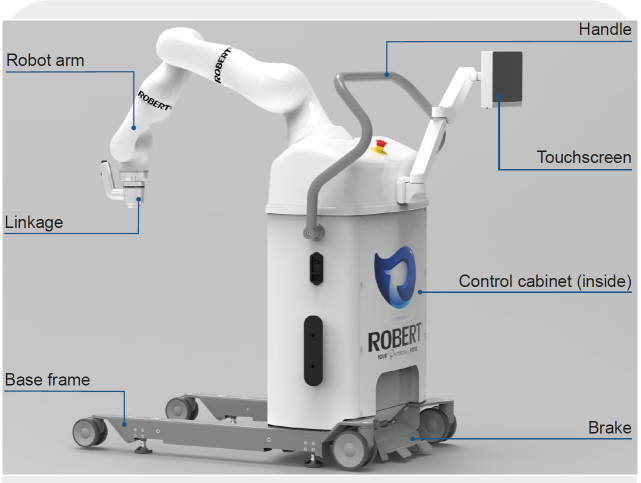


Fig. 1. ROBERT® system overview.

There is no built-in program in the robot. It only replicates the exercise path performed by PT. The robot is a tireless workhorse to allow up to 999 repetitions for one exercise. Four exercises can be recorded in one go. The quality of exercise is the same as PT. Repetitive exercises are beneficial to neuroplasticity rewiring. After the PT performs the first cycle of exercise, and hereafter, the robot will start replicating the exercise in certain cycles without any support from the PT. Then, the PT can treat another patient or do some documentation work. The PT efficiency is improved.

Ho *et al.* mentioned that rehabilitation process should start as early as possible to avoid spasticity.[14] The novel robot is good for many stages of lower limb rehabilitation. It allows passive or active mode. Patients in poor condition can be trained in passive mode that the robot brings the impaired leg to move. After the leg gains power, patients can be trained in active mode. The starting point and ending point are set. With no restriction on path, the patient can move his/her leg to reach the destination autonomously, which is good for motor control and learning. There are 10 levels of resistive force for active training. A healthy adult feels hard when the resistive force reaches level 7 or above. The resistive force and number of cycles from the active exercise can form a treatment evolution graph. Thus, a better individualized treatment can be prepared. The robot allows treatment in supine, prone, lateral or sitting positions. This serves bedridden or wheelchair patients. Furthermore, the Danish robot is complementary to exoskeleton devices. Prior to being trained with a walking device, patients can be trained with ROBERT® for hip flexion to attain a large extent and for dorsal/plantar orientations (Dorsiflexion 20° — Plantar flexion 70°) to strengthen their ankles. These improvements would make a walking exercise session more efficient.

The robot is equipped with emergency buttons: One was installed on the control cabinet and one is held by the patient. For safety precaution, the device will automatically stop moving if the patient's leg is against the movement of the robotic arm, e.g., spasticity or cramp.

The device is mobile with four wheels and weighs 165 kg. It can be moved between wards or stayed in outpatient department. It is equipped with a built-in battery of 30-min power capacity to enable the robot to be powered on during transportation. This saves users four minutes for system

rebooting and self-calibration. The device is ready for treatment when it arrives at the bedside of the next patient.

The setup interfaces for PTs are intuitive. That leads the setup quick and learning curve short. A 10-inch touch screen is installed on a swingable holder. During training, the screen is swung close to the patient. The software can be displayed in Chinese or some other popular European languages and the visual is aided by graph presentation. Patients are able to understand the training progress easily.

Bertelsen *et al.* gave feedback in their pilot test of using ROBERT® that about 10–15 min were spent on setup including moving furniture to make space in the ward and attaching the velcro-sheet fixture to the training leg.[13] This problem will happen in Hong Kong as well. However, after the setup, patients can be trained uninterruptedly for the rest of the session.

Although other robotic therapy machines are capable to perform the exercises robotically, a PT is required to operate them. There is no manpower saving. However, the ROBERT® frees up the PT to help another patient or do other work after the setup. It truly optimizes PT's precious time.

Cost Effectiveness Analysis

Here is a brief cost-effectiveness analysis between ROBERT® and PT and its details are shown in Table 1.

Cost: (1) The robot: The reference price of ROBERT® in June 2022 was HK$1,509,200.00 with 10,000 h expected lifetime. The average hourly operating cost is **HK$175.92**; whereas

(2) The PT: 2022 salary report of Physiotherapists in Hong Kong stated that a mean salary of PT is HK$63,000.00.[15] Thus, the hourly salary of PT is **HK$269.23**.

Effectiveness: The effectiveness of ROBERT® and PT is the same because the robot replicates what PT performs. Either effectiveness result is defined as "one eff. unit".

Cost-Effectiveness Analysis: Dividing the Cost by the Effectiveness for each group, the cost per one eff. unit is calculated.

$$\text{Cost Effectiveness} = \frac{\text{Cost}}{\text{Effectiveness}}.$$

A larger value in cost-effectiveness comparison indicates higher cost be paid to obtain the same result.

The cost-effectiveness analysis result showed that employing Danish robot to perform repetitive exercises for lower limbs is less costly than the equivalent performed by PT. Its cost-effectiveness is **HK$175.92**/one eff. unit; whereas that of PT is **HK$269.23**/one eff. unit. Although the capital cost of the robotic system is high, its average hourly operating cost is just one-tenth of the cost for one specialty outpatient session in HA hospitals.[7]

Table 1. Cost-effectiveness analysis among ROBERT® and PT.

(All currencies are in HK$)	ROBERT®	Physiotherapist
COST		
a. Therapist salary	—	$63,000
b. Daily working hours	5.25	9.00
c. Monthly working days	26	26
d. Machine cost	$1,509,200.00	—
e. Annual maintenance fee	$50,000.00	—
f. Number of year with annual maintenance [*Expected Operating Hour/(b x c x 12)hour/year*] — *Warranty years*	5 *(2-year warranty)*	—
g. Total machine cost [$d + (exf)$]	$1,759,200.00	—
h. Total hourly cost	$175.92	$269.23
EFFECTIVENESS	one eff. unit	one eff. unit
RESULT		
Cost effectiveness	**$175.92**/eff. unit	**$269.23**/eff. unit

Discussion

Due to a shortage of clinicians, the Hong Kong healthcare system unavoidably considers robots to assist in maintaining or enhancing service quality. The healthcare system is progressing into the robotic era. A Robot-Assisted Gait Training device — Lokomat — was awarded in April 2022 for Kowloon East Cluster. There are a wide variety of evaluations of its effectiveness in locomotor training. van Nunen *et al.* concluded that there was no significant difference between Lokomat and the conventional training,[16] while Duncan *et al.* showed a better effectiveness ratio of 2.33 than PT did.[17] Based on the result from Duncan *et al.* and the amount of HK$ 1,950,000 on the contract award notice webpage, the cost-effectiveness of Lokomat was translated to about HK$210.55/one eff. unit in locomotor training versus that of PT is HK$269.23/one eff. unit manually.

Robotic treatment with ROBERT® fulfils Lum *et al.*'s criteria of choosing a Robot therapy system. It clinically benefits the patients, especially stroke patients. Therapists' efficiency will be improved. Its average cost is less than HK$200 per operating hour. It potentially helps reduce the economic burden by reducing LOS and lowering the readmission rate.

Updated on the end of 2021, there have been 13 systems in clinical use, mainly in neurology centres in Denmark and Germany. This technology is innovative and swiftly accepted. Throughout the world, hundreds of patients are treated with it on a daily basis.

Conclusion

The Danish robot benefits patients, PT, and HA quantitatively and qualitatively. It improves patients' condition, increases patient numbers, releases PT shortage stress, shortens bed occupancy time, and reduces the economic burden. That the robot performs repetitive exercises for lower limbs is less costly than the equivalent performed by PT. ROBERT® potentially would be another cost-effective solution for HA.

Conflict of Interest

The author is an employee of a distributor company of ROBERT®.

Author Contributions

The author conducted the conception and design of the study, drafted this paper and performed revision of the final version of this paper.

References

1. Daras LC, Deutsch A, Ingber MJ, Hefele JG, Perloff J. Inpatient rehabilitation facilities' hospital readmission rates for medicare beneficiaries treated following a stroke. Top Stroke Rehabil 2021;28(1):61–71.
2. Hancock N, Kilbride C, McGregor L. Physiotherapy concise guide for stroke 2016. Royal College of Physicians, United Kingdom, 2016, www.strokeaudit.org/guideline.
3. Slade A, Tennant A, Chamberlain A. A randomised controlled trial to determine the effect of intensity of therapy upon length of stay in a neurological rehabilitation setting. J Rehabil Med 2002;34:260–66.
4. Woo J, Ho SC, Goggins W, Chau PH, Lo SV. Stroke incidence and mortality trends in Hong Kong: Implications for public health education efforts and health resource utilization. Hong Kong Med J 2014;20(S3):S24–9.
5. Chau PH, Woo J, Goggins W, Tse Y, Chan K, Lo SV, et al. Trends in stroke incidence in Hong Kong differ by stroke subtype. Cerebrovasc Dis 2011;31:138–46.
6. Fung ML. Stroke Rehabilitation: Predicting inpatient length of stay and discharge placement. Hong Kong J Occup Ther 2004;14:3–11.
7. Hospital Authority Annual Report 2020-2021 Appendixes 8 & 11. Available at https://www.ha.org.hk/haho/ho/cc/HA_Annual_Report_2020-21_en.pdf.
8. Hospital Authority Statistical Report 2020-2021; Cluster Reports. Available at https://www3.ha.org.hk/data/HAStatistics/StatisticalReport/2020-2021.
9. Shortage of physios sparks warning. The Standard, 19 Jun 2017. Available at https://www.thestandard.com.hk/section-news/section/4/184119/Shortage-of-physios-sparks-.warning#:~:text=A%20manpower%20shortage%20projection%20on%20the%20physiotherapy%20sector,see%20a%20shortage%20of%20933%20physiotherapists%20by%202030.
10. Kwok T, Ma E. Clinical applications of technology in rehabilitation. The Hong Kong Med Dairy 2013;18(7):21–2.
11. Lum P, Reinkensmeyer D, Mahoney R, Rymer WZ, Burgar C. Robotic devices for movement therapy after stroke: Current status and challenges to

clinical acceptance. Top Stroke Rehabil 2002;8(4):40–53.
12. Hug A, Spingler T, Hensel C, Fichtner S, Daniel T, Heutehaus L, et al. Goal attainment in mobility after acute rehabilitation of mobility-restricting paralysis syndromes with regard to the ambulatory therapeutic level of participation NeuroMoves. A German national multicenter observational cohort study. BMC Neurol 2021;21(149):1–13.
13. Bertelsen AS, Storm A, Minet L, Ryg J. Use of robot technology in passive mobilization of acute hospitalized geriatric medicine patients: A pilot test and feasibility study. Pilot Feasibility Stud 2020;6(1):1–9.
14. Ho L, Ngan PS. Spasticity. News Bull. Hong Kong Physiother. Assoc 2019;23(4):1.
15. Average Salary of Hong Kong Physiotherapists at year 2022. Available at http://www.salaryexplorer.com/salary-survey.php?loc=97&loctype=1&job=916&jobtype=3.
16. van Nunen MPM, Gerrits KHL, Konijnenbelt M, Janssen TWJ, de Haan A. Recovery of walking ability using a robotic device in subacute stroke patients: A randomized controlled study. Disabil Rehabil Assist Technol 2014;10(2):141–8.
17. Duncan PW, Sullivan KJ, Behrman AL, Azen SP, Wu SS, Nadeau SE, Dobkin BH, Rose DK, Tilson JK. Protocol for the locomotor experience applied post-stroke (LEAPS) trial: A randomized controlled trial. BMC Neurol 2007;7:39–62.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 42, No. 2 (2022) 81–89
DOI: 10.1142/S101370252250007X

Hong Kong Physiotherapy Journal
https://www.worldscientific.com/worldscinet/hkpj

Effectiveness of Acapella along with institutional based chest physiotherapy techniques on pulmonary functions and airway clearance in post-operative CABG patients

Bhagyashree Jage and Anuprita Thakur*

School of Physiotherapy, D. Y. Patil University
Navi Mumbai, Maharashtra, India

**anupritathakur@yahoo.com*

Received 27 May 2021; Accepted 9 February 2022; Published 15 June 2022

Background: Patients undergoing Coronary Artery Bypass Graft (CABG) surgery often develop pulmonary complications in the early post-operative period as result of decreased lung function and impaired cough. Conventional physiotherapy in early post-operative period aims at increasing lung volumes and airway clearance.
Objective: This study aimed to determine the effectiveness of the addition of Acapella to conventional chest physiotherapy in improving lung volumes and secretion clearance in early post-operative CABG patients.
Methods: Twenty patients of both genders (40–70 years) who had undergone CABG and were in Phase I of Cardiac Rehabilitation were involved in this pilot randomized control trial (9 control, 11 experimental). Post-surgery intervention commenced on post-operative day 2 (POD 2) and continued till POD 6. Patients in the control group were given conventional physiotherapy that included breathing exercises, incentive spirometry and manual techniques. Patient in the experimental group used an Acapella device along with the conventional intervention. Outcome measures considered were pulmonary function parameters (FVC, FEV1 & PEFR) and amount of sputum expectorated.
Results: A significant increase in lung volumes was observed in both the groups on POD 6 as compared to POD 2 (both < 0.01). However, the increase was significantly greater on POD 6 in experimental group than the control group [mean difference (95% CI) FVC: 0.44 L (0.24–0.63), FEV1: 0.43 L (0.19–0.66), PEFR: 0.86 L/s (0.57–1.14)]. The amount of sputum expectoration significantly greater in the experimental group as compared to the control group [2.71 mL (0.53–4.90)].

*Corresponding author.

Copyright@2022, Hong Kong Physiotherapy Association. This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) License which permits use, distribution and reproduction, provided that the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Conclusion: The addition of Acapella enhanced the effect conventional physiotherapy in improving lung volumes and airway clearance in the early post-operative period for CABG patients.

Keywords: Acapella; CABG; chest physiotherapy.

Introduction

Patients undergoing cardiac surgeries are prone to distinct surgery related factors that influence the development of certain post-operative pulmonary complications (PPCs).[1] Studies reveal a post-operative change in breathing pattern from predominantly abdominal to thoracic breathing resulting in reduction of pulmonary functions in patients undergoing median sternotomy.[2,3] A decrease in VC, IC, FEV1, PEFR, TLC leading to a restrictive pattern of pulmonary function is also observed post cardiac surgery.[4–8] The causes cited for this decrease are multifactorial; sternotomy incision leading to decreased rib cage movement, respiratory depression due to anesthesia, diaphragm dysfunction leading to alterations in breathing patterns, presence of chest tubes, post-operative pain.[1–3,6,9–11] This significant impairment in pulmonary function is observed to persist up to almost one year post-operatively.[2,4–7,9] Additionally post-operative atelectasis is also observed in CABG patients,[7,12] which can further worsen the pulmonary function. A study done by Koyilil *et al.*[13] reported increased incidence of cough post-operatively following open-heart surgery. The severity of the cough correlated with the reduction in lung volume and was independent of the basic cardiac pathology and left ventricular function, cardiopulmonary bypass duration and smoking status. The researchers concluded that the primary cause of cough was the post-operative reduction in lung volume. While the post-operative decrease in PEFR reduces the ability to cough, the general anesthesia used during surgery causes impairment of mucociliary transport, both resulting in sputum retention.[6,14,15]

To summarize, the change in breathing pattern, reduction in lung volumes and sputum retention contribute to post-operative complications in CABG patients. A study done by Agostini *et al.* reported that patients with PPCs had a significantly higher hospital length of stay and higher frequency of ICU admissions and number of deaths.[16] This puts an additional burden on the patient and hospital.

In order to decrease the after-effects of decreased lung volumes and impaired secretion clearance post-operatively, physiotherapy interventions in the form of breathing exercises, incentive spirometry and physical maneuvers to recruit alveoli are often recommended.[17,18] Manual techniques, such as percussions and vibrations, are commonly used to aid removal of secretions.[19] However in the post-operative period, these techniques may not be well tolerated by all patients due to post-operative pain[20] and may be contraindicated in some patients due to inability to change body position. Some patients may not be able to self-apply these techniques without assistance and thus adherence and regular administration may depend on availability of therapist or caregiver. In addition, these techniques are also laborious and time consuming for the therapist.[20] Hence there is an increase in the use of assistive devices to aid removal of secretions. In the recent years, various devices have emerged that assist existing physiotherapy techniques in improving mobilization and removal of secretions.[21] These devices are safe, offer greater independence to patients and are less time-consuming for the therapist. Acapella is one such flow-operated Oscillatory Positive Expiratory Pressure (OPEP) device that consists of a counterweighted plug and metal strip that is attached to a magnet which oscillates when the patient exhales into the device resulting in airflow oscillations.[22] The resulting positive expiratory pressure (PEP) and oscillations assist mucus expectoration.[19,21]

Studies regarding efficacy of Acapella as an aid to clear secretions has been evaluated in various pulmonary conditions,[23–25] however its scope to improve pulmonary function and cough mechanism in post cardiac surgery patients is insufficiently researched. This study hypothesized that the addition of Acapella to conventional Institutional Chest physiotherapy would aid in improving lung volumes and sputum expectoration compared to conventional Institutional Chest physiotherapy

alone in patients undergoing CABG surgery in the early post-operative period.

Methodology

This was a single-site randomized controlled feasibility pilot trial with assessor blinding and intention-to-treat analysis.

Approval to conduct this study was obtained from the Institutional Review Board, D. Y. Patil University, School of physiotherapy (DYPUSOP/019A/2018).

Participants

Patients of both genders in the age group of 40–70 years who had undergone CABG (through median sternotomy incision) and were currently hospitalized in the ICU of D.Y. Patil Hospital and Research Centre, Navi Mumbai, India and undergoing Phase 1 of cardiac rehabilitation were selected for this study. The number of patients selected in the study was based on the patient availability during the study period. Oral informed consent was taken from all the participating patients. Any patient with existing pulmonary disease or associated pulmonary complications, unstable cardiovascular status, infection, sepsis, uncontrolled diabetes, other metabolic problems and impaired cognition were excluded from the study. Additionally, patients who required long-term intubation post-operatively or reported too much pain (in spite of medications) were not included in the study. The patients were randomly allocated to either control or experimental groups using the lottery method – paper slips with numbers written and folded were handed to the participants. Patients who picked an odd number were allocated to the control group and an even number were allocated to the experimental group. After group allocation, the assessment and intervention procedures were explained to all participants.

Outcome measures

The primary outcome measures considered in the study were lung volumes measured using PFT (FVC, FEV1, PEFR) and the cumulative amount of sputum expectorated by Post-operative day 6 (POD 6).

The assessment began on POD 2 and was same for both the groups. Pain was assessed on the Numerical Rating Scale before intervention. The pulmonary function test (PFT) parameters were assessed pre- and post-intervention on POD 2 and post-intervention on POD 4 and POD 6. A handheld portable spirometer (EasyOne^{TM} Diagnostic Spirometer, ndd Medizintechnik AG, Switzerland) was used to measure PFT by the therapist. For this, patient was seated upright with back supported. The mouthpiece of spirometer was placed into patient's mouth and was asked to form a tight seal around it. Patient was then instructed to inspire at total lung volume and nose clip was applied. He/she then exhaled forcefully into the mouthpiece. Three attempts for FVC, FEV1 and PEFR were recorded and the highest value from

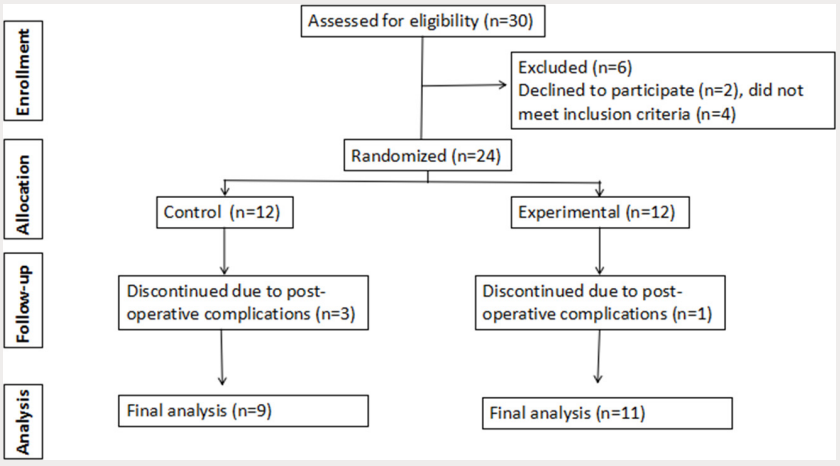


Fig. 1. Flowchart showing participation through the study.

the best attempt for each parameter was used in the data analysis.

Amount of sputum expectorated during every intervention (for each session/day) was recorded by collecting it in a sputum container with markings on it. The total amount of sputum expectorated (in mL) from POD 2 to POD 6 was noted on POD 6.

Interventions

Similar post-operative medical care in terms of ventilation and pain medications and standard institutional cardiac rehabilitation protocol (e.g., bedside mobilizations and ambulation) was followed for all the patients as per the hospital policy. Chest physiotherapy intervention with respect to the study, started on POD 2 for both groups. The same therapist (principal investigator of this study) administered the interventions in both groups for all sessions. The average time for treatment ranged between 20 min and 40 min. The interventions were given with the patients lying with back supported and the head end angled at 45°.[26]

For the control group, the treatment included diaphragmatic breathing exercises[19] (10 repetitions, 2 sets, twice a day) and segmental breathing exercises (10 repetitions, 3 sets, twice a day) along with percussions and vibrations. Incentive spirometry[19] was given as 10 repetitions of 3 sets, twice a day. Patient was encouraged to cough out and expectorate secretions, if any, in a measured cup.

In the experimental group, similar intervention as above was followed, except instead of percussion and vibration, Acapella green[27,28] (Smiths Medical ASD, Inc., USA) was prescribed twice a day with at least 4 hourly gaps between two sessions. For Acapella, the patients were asked to hold the device with a tight seal at the mouth, take a breath larger than normal tidal volume and hold for 2–3 s before exhaling into the device.[19] This exhalation into Acapella was repeated 8–10 times after which the patients were instructed to perform 2–3 huffs and expectorate in the measuring cup. This cycle of 8–10 breaths in Acapella followed by 2–3 huffs was repeated 4–6 times. This set was performed twice a day.[29]

The study ceased on POD 6 as the average stay of post CABG patients in our ICU ranged between 6 and 8 days.

Statistical analysis

The data was analyzed using SPSS software (version 18). For any within-group comparison of outcomes, one-way ANOVA was used with Bonferroni as the post-hoc test. For any between-group comparisons of outcomes (i.e., between control and experimental groups), an unpaired '*t*' test was performed. The level of significance was set at < 0.005.

Results

A total of 30 subjects were considered for inclusion in the study, of which 6 had to be excluded. The remaining 24 subjects were randomly allocated to control and experimental group. None of the included subjects had any major co-existing co-morbidity except diabetes. Four patients (3 from control and 1 from experimental group) dropped out due to post-operative complications (infection at the scar site and arrhythmia), while 20 patients completed the study procedures. The general demographics of the subjects included in the study are represented in Table 1. The mean age and BMI of the four subjects who dropped out was 51.25 ± 5.737 and 26.36 ± 4.478.

All the patients had undergone CABG on pump. Of the 20 subjects, 10 patients had only saphenous vein graft, 4 had only internal mammary artery graft and 6 had arterial and venous grafts.

The pulmonary volumes (FVC, FEV1 and PEFR) and the total amount of sputum production in both groups showed a statistical improvement post intervention on all the days. However, it was observed that the improvement was more on POD6 in both the groups (Table 2). Table 3 compares the mean differences in the lung volumes between POD

Table 1. Descriptive demographics.

Variables	Control ($n = 12$)	Experimental ($n = 12$)	p value
Males/Females	8/4	11/1	
Age (in years)	57.75 ± 4.95	54.5 ± 5.23	0.193
BMI	24.64 ± 2.66	23.77 ± 1.66	0.295
Smokers	2	3	
Controlled diabetics	2	3	

Note: Data presented is as mean ± SD BMI: Body Mass Index.

Table 2. Comparison of lung volumes and amount of sputum expectorated within and between control and experimental groups.

	Control	Experimental	*P* value (between experimental and control)	Mean differences (95% CI) (between experimental and control)
FVC (in L)				
Pre intervention POD 2	0.26 ± 0.06	0.39 ± 0.13	**0.020**	**0.12 (0.02–0.22)**
Post intervention POD 2	0.32 ± 0.05	0.55 ± 0.31	**0.043**	**0.23 (0.007–0.46)**
Post intervention POD4	0.44 ± 0.10	1.08 ± 1.25	**0.014**	**0.64 (0.24–1.52)**
Post intervention POD 6	0.54 ± 0.12	1.11 ± 0.40	**< 0.001**	**0.56 (0.27–0.86)**
P value (from Pre intervention POD 2 to POD 6)	**< 0.001**	**0.001**		
FEV1 (in L)				
Pre intervention POD 2	0.25 ± 0.06	0.42 ± 0.23	**0.057**	**0.16 (0.005–0.34)**
Post intervention POD 2	0.30 ± 0.04	0.51 ± 0.20	**0.009**	**0.20 (0.05–0.35)**
Post intervention POD4	0.4033 ± 0.10	0.72 ± 0.27	**0.004**	**0.32 (0.11–0.53)**
Post intervention POD 6	0.51 ± 0.10	1.11 ± 0.48	**0.002**	**0.60 (0.25–0.94)**
P value (from Pre intervention POD 2 to POD 6)	**< 0.001**	**< 0.001**		
PEFR (in L/s)				
Pre intervention POD 2	0.57 ± 0.12	0.79 ± 0.36	**0.112**	**0.21 (−0.05 to 0.48)**
Post intervention POD 2	0.63 ± 0.13	1.04 ± 0.37	**0.006**	**0.41 (0.13–0.68)**
Post intervention POD4	0.87 ± 0.39	1.47 ± 0.47	**0.007**	**0.60 (0.19–1.02)**
Post intervention POD 6	0.94 ± 0.17	2.02 ± 0.49	**< 0.001**	**1.07 (0.71–1.44)**
P value (from Pre intervention POD 2 to POD 6)	**0.006**	**< 0.001**		
Total Sputum production on POD 6 (in mL)	7.55 ± 1.01	10.27 ± 2.96	**0.018**	**2.71 (0.53–4.90)**

Notes: Data presented as mean ± SD (95% CI). POD – post-operative day, FVC – Forced Vital Capacity, FEV1 – Forced Expiratory Volume in first second, PEFR – Peak Expiratory Flow Rate, POD – Post-operative day, CI – confidence interval.

6 and pre intervention POD 2 values for both the groups. A 107% and 104% increase, respectively, in FVC and FEV1 from POD 2 pre-intervention to POD 6 was observed in the control group while the experimental group showed an increase by 184% and 164%, respectively, for the same parameters. PEFR increased by 64% from POD 2 pre-intervention to POD 6 in the control group, while a 155% increase was observed in the experimental group. The results indicate that the experimental group showed a significantly greater improvement in lung volumes on POD 6 as compared to the control group. The sputum production was also significantly more in the experimental group.

Discussion

The results of this study indicate that inclusion of Acapella along with other conventional physiotherapy techniques increased the lung volumes and airway clearance as compared to conventional physiotherapy techniques alone in post-operative CABG patients.

Table 3. Comparison of the mean differences in lung volumes between POD 6 and POD 2 Pre-intervention between the experimental and control group.

	Control	Experimental	*P* value (between mean differences of POD 6 and Pre intervention POD 2 between experimental and and control)	Mean differences (95% CI) (POD 6 and Pre intervention POD 2 between experimental and control)
FVC (in L)	0.28 ± 0.09	0.72 ± 0.26	**0.0002**	**0.44**
[Mean differences between POD 6 and Pre intervention POD 2 (95% CI)]	(0.18–0.37)	(0.45–0.98)		**(0.24–0.63)**
FEV1 (in L)	0.26 ± 0.08	0.69 ± 0.33	**0.0013**	**0.43**
[Mean differences between POD 6 and Pre intervention POD 2 (95% CI)]	(0.17–0.34)	(0.35–1.02)		**(0.19–0.66)**
PEFR (in L/s)	0.37 ± 0.14	1.23 ± 0.38	**< 0.001**	**0.86**
[Mean differences between POD 6 and Pre intervention POD 2 (95% CI)]	(0.22–0.51)	(0.84–1.16)		**(0.57–1.14)**

Notes: Data presented as mean ± SD (95% CI). POD – post-operative day, FVC – Forced Vital Capacity, FEV1 – Forced Expiratory Volume in first second, PEFR – Peak Expiratory Flow Rate, POD – Post-operative day, CI – confidence interval.

Acapella is shown to be effective in aiding sputum clearance in variety of conditions.[25,30] Studies in mechanically ventilated ARDS patients showed that Acapella aided in optimally clearing the secretions from the airways.[23,31] Studies in bronchiectasis patients also demonstrated an increase in sputum volume production following use of Acapella.[24,32]

Acapella combines the resistive effect of PEP with high-frequency oscillations in the airways during exhalation to facilitate secretion clearance.[27,33–35] PEP allows back pressure to be generated that opens and splints the peripheral airways. This encourages collateral ventilation and airflow to move behind the secretions.[23,34–37] The pressure gradient across the secretions forces it to move more centrally and thus help in secretion clearance.[36,37] In addition, the oscillations produce vibrations within the airway wall that further help to displace secretions into the airway lumen.[33–35] Some studies indicate that the oscillations generated by the OPEP devices can cause break down of the mucus macro-molecules bonds reducing the viscoelasticity (thickness) of the secretions and thus further enhance their transport through the airways.[38,39]

Another study that assessed the pressure characteristics of Acapella under laboratory conditions, suggested that the oscillation frequency range produced by Acapella (8.5–21 Hz)[33,40] coincides with that of ciliary beating frequency (12–15 Hz) in tracheobrochial tree[41,42] and respiratory system resonance frequency[33] and this also facilitates the secretion movement.[43] Comparative studies between Acapella and other PEP devices (Flutter, Shakers) reveal that though these devices have similar operating performances and produce similar pressure waveforms, Acapella produces higher frequency of oscillation than those of other PEP devices at low-pressure levels.[27,44]

In this study, we too observed an increase in sputum production in the experimental group using Acapella as compared to the control group, probably through the mechanisms explained earlier. This was also accompanied with an increase in lung volumes, though earlier studies show mixed results. Use of Acapella did not change lung function following lung resection surgery[20] or bronchiectasis,[24] however a significant improvement in lung volumes was observed in patients who had undergone upper abdominal surgeries,[45] video-assisted thoracic surgery[46] and CABG.[47] The change in lung volumes observed in this study could be a cumulative effect of breathing exercises[19,30] and better airway clearance due to Acapella.

Moreover, patients reported to be comfortable with the use of Acapella as compared to incentive spirometry[20,32] or manual techniques like percussion and vibration which probably was because Acapella did not irritate or stimulate the chest wall or wound directly but internally transmitted the vibration to the secretions in the airways.[20] As

Acapella is not gravity-dependent, it is easier to use in patients with low expiratory flow rates and in whom change of positions is difficult.[27,44]

The other advantages of Acapella lie in the fact that it is available in different models that allow selection of device based on the patient's expiratory flow capacity. Furthermore, Acapella can be used along with nebulizer[21] whenever the need be. All the above-mentioned advantages of Acapella may have helped the patients in improving their lung volumes and aided in easier clearance of secretions in subjects in this study.

This study has some limitations. The sample size in the study was small as it was confined by the number of patients available at the time of study. To establish the clinical relevance of the percent changes observed in the lung volumes between pre and post intervention, future studies with larger sample sizes is recommended. The pre-operative spirometry values were also not available, hence the immediate post-operative change in lung volumes could not be assessed. In addition, the study was restricted to POD6 for reasons mentioned earlier. Future studies for longer periods of time to assess the long term effects are suggested for generalizing the usability of Acapella in patients undergoing CABG.

Conclusion

This pilot RCT suggests that the effect of conventional physiotherapy techniques in improving lung volumes and secretions clearance is enhanced with the use of Acapella in the post-operative period in CABG patients. However, further studies with larger sample size and patients with co-existing pulmonary conditions are recommended to establish the effectiveness of Acapella in post-cardiac surgery patients. Studies of longer duration can be conducted to assess the long-term effect of Acapella in post-operative cardiac patients.

Conflict of Interest

The authors declare that there is no conflict of interest.

Funding/Support

Not applicable.

Author's Contributions

The study was conceptualized and designed by BJ and AT. BJ did the data collection and drafted the initial paper. AT critically reviewed the paper for intellectual content and subsequently revised the paper. Both the authors approved the final version of the paper.

References

1. Wynne R, Botti M. Postoperative pulmonary dysfunction in adults after cardiac surgery with cardiopulmonary bypass: Clinical significance and implications for practice. Am J Crit Care 2004;13:384–93.
2. Ragnarsdottir M, KristjAnsdottir A, Ingvarsdottir I, Hannesson P, Torfason B, Cahalin L. Short-term changes in pulmonary function and respiratory movements after cardiac surgery via median sternotomy. Scand Cardiovasc J 2004;38:46–52.
3. Roncada G, Dendale P, Kinsen L, Hendrikx M, Hansen D. Reduction in pulmonary function after CABG surgery is related to postoperative inflammation and hypercortisolemia. Int J Clin Exp Med 2015;8:10938–46.
4. Shenkman Z, Shir Y, Weiss YG, Bleiberg B, Gross D. The effects of cardiac surgery on early and late pulmonary functions. Acta Anaesthesiol Scand 1997;41:1193–99.
5. Johnson D, Hurst T, Thomson D, Mycyk T, Burbridge B, To T, et al. Respiratory function after cardiac surgery. J Cardiothorac Vasc Anesth 1996;10:571–7.
6. Westerdahl E, Lindmark B, Bryngelsson I, Tenling A. Pulmonary function 4 months after coronary artery bypass graft surgery. Respir Med 2003;97:317–22.
7. Morsch KT, Leguisamo CP, Camargo MD, Coronel CC, Mattos W, Ortiz LD, et al.. Ventilatory profile of patients undergoing CABG surgery. Braz J Cardiovasc Surg 2009;24:180–7.
8. Goyal V, Pinto RJ, Mukherjee K, Trivedi A, Sharma S, Bhattacharya S. Alteration in pulmonary mechanics after Coronary Artery Bypass Graft surgery: Comparison using internal mammary artery and saphenous vein grafts. Indian Heart J 1994;46:345–8.
9. Baumgarten MC, Garcia GK, Frantzeski MH, Giacomazzi CM, Lagni VB, Dias AS, et al.. Pain and pulmonary function in patients submitted to heart surgery via sternotomy. Braz J Cardiovasc Surg 2009;24:497–5.
10. Large SR, Heywood LJ, Flower CD, Cory-Pearce R, Wallwork J, English TA. Incidence and etiology

of a raised hemidiaphragm after cardiopulmonary bypass. Thorax 1985;40:444–47.
11. Siegel W, Loop FD. Comparison of internal mammary artery and saphenous vein bypass grafts for myocardial revascularization: Exercise test and angiographic correlations. Circulation 1976;54: III1–III3.
12. Yánez-Brage I, Pita-Fernández S, Juffé-Stein A, Martínez-González U, Pértega-Díaz S, MauleónGarcía A. Respiratory physiotherapy and incidence of pulmonary complications in off-pump coronary artery bypass graft surgery: An observational follow-up study. BMC Pulm Med 2009;9:36.
13. Koyilil v, Mohanakrishnan, Rieyaz H, Rajpathy J, Selvakumar P, Deneshkumar V. Non infective cough in immediate post cardiac surgery patients. Int J Contemp Med Surgery Radiol. 2019;4:D75–D80.
14. Keller C, Brimacombe J. Bronchial mucus transport velocity in paralyzed anesthetized patients: A comparison of the laryngeal mask airway and cuffed tracheal tube. Anesth Analg 1998;86:1280–82.
15. Miskovic A, Lumb AB. Postoperative pulmonary complications. BJA: Brit J Anaesthesia 2017;118:317–34.
16. Agostini P, Cieslik H, Rathinam S, Bishay E, Kalkat MS, Rajesh PB, Steyn RS, Singh S, Naidu B. Postoperative pulmonary complications following thoracic surgery: Are there any modifiable risk factors? Thorax 2010;65:815–18.
17. Westerdahl E, Lindmark B, Eriksson T, Friberg O, Hedenstierna G, Tenling A. Deep-breathing exercises reduce atelectasis and improve pulmonary function after coronary artery bypass surgery. Chest 2005;128:3482–88.
18. Cavenaghi S, Ferreire LL, Marino LH, Lamari NM. Respiratory physiotherapy in the pre and postoperative myocardial revascularization surgery. Braz J Cardiovasc Surg 2011;26:455–61.
19. Pryor J, Prasad A. Physiotherapy for Respiratory and Cardiac Problems: Adults and Pediatrics (Physiotherapy Essentials). 4th edition. Churchill Livingstone (Elsevier Ltd); 2013. Chap 5.
20. Cho YJ, Ryu H, Lee J, Park IK, Kim YT, Lee YH, et al.. A randomised controlled trial comparing incentive spirometry with the Acapella device for physiotherapy after thoracoscopic lung resection surgery. Anaesthesia 2014;69:891–98.
21. Hristara-Papadopoulou A, Tsanakas J, Diomou G, Papadopoulou O. Current devices of respiratory physiotherapy. Hippokratia 2008;12:211–20.
22. Patterson JF, Bradley JM, Hewitt O, Bradbury I, Elborn JS. Airway clearance in bronchiectasis: A randomized crossover trial of Active Cycle of Breathing techniques versus Acapella. Respiration 2005;72:239–42.
23. Sharma P, Prem V, Jain S. Immediate effects of Acapellaon dynamic lung compliance in mechanically ventilated patients with acute respiratory distress syndrome: A case series. Indian J Crit Care Med 2018;22:100–2.
24. Patterson JE, Hewitt O, Kent L, Bradbury I, Elborn JS, Bradley JM. Acapella versus 'usual airway clearance' during acute exacerbation in brochiectasis: A randomized crossover trial. Chron Respir Dis 2007;4:67–74.
25. Rodrigues A, Castro GM, Jacome C, Langer D, Parry SM, Burtin C. Current developments and future directions in respiratory physiotherapy. Eur Respir Review 2020;29:200264.
26. Guner SI, Korkmaz FD. Investigation of the effects of chest physiotherapy in different positions on the heart and the respiratory system after coronary artery bypass surgery. Toxicol Ind Health 2015;31:630–37.
27. Volsko TA, DiFiore J, Chatburn RL. Performance comparison of two oscillating positive expiratory pressure devices: Acapella versus Flutter. Respir Care 2003;48:124–30.
28. Mueller G, Bersch-Porada I, Koch-Borner S, Raab A, Jonker M, Bamberger M, Michel F. Laboratory evaluation of four different devices for secretion mobilization: Acapella Choice, Green and Blue versus water bottle. Respir Care 2014;59:673–77.
29. Ahmad A. Essentials of physiotherapy after Thoracic surgery: What Physiotherapists need to know. A Narrative Review. Korean J Thorac Cardiovasc Surg 2018;51:293–307.
30. Bott J, Blumenthal S, Buxton M, Ellum S, Falconer C, Garrod R, et al.. British thoracic society physiotherapy guideline development group. Guidelines for the physiotherapy management of adult, medical, spontaneously breathing patient. Thorax 2009;64:i1–i51.
31. Narula D, Nangia V. Use of an oscillatory PEP device to enhance bronchial hygiene in a patient of post-H1NI pneumonia and acute respiratory distress syndrome with pneumothorax. BMJ Case Rep 2014;7:2014.
32. Naraparaju S, Vaishali K, Venkatesan P, Acharya V. A comparison of the Acapella and a threshold inspiratory muscle trainer for sputum clearance in bronchiectasis-A pilot study. Physiother Theory Pract 2010;26:353–57.
33. Silva C, Santos J, Jansen j, Lopes de Melo P. Laboratory evaluation of the Acapella device: Pressure characteristics under different conditions, and a software tool to optimize its practical use. Respir Care 2009;54:1480–87.

34. App E, Kieselmann R, Reinhardt D, Lindemann H, Dasgupta B, King M, et al. Sputum rheology changes in cystic fibrosis lung disease following two different types of physiotherapy. Flutter vs autogenic drainage. Chest 1998;114:171–77.
35. Tambascio J, De Souza L, Lisboa R, Passarelli R, Souza H, Gastaldi A. The influence of flutterVRP1 components on mucus transport of patients with bronchiectasis. Respir Med 2011;105:1316–21.
36. Falk M, Kelstrup M, Andersen J, Kinoshita T, Falk P, Støvring S, Gothgen I. Improving the ketchup bottle method with positive expiratory pressure, PEP, in cystic fibrosis. Eur J Respir Dis 1984;65:423–32.
37. Van der Schans C, Van der Mark T, De Vries G, Piers D, Beekhuis H, Dankert-Roelse J, Postma D, Koeter G. Effect of positive expiratory pressure breathing in patients with cystic fibrosis. Thorax Apr 1991;46:252–6.
38. Altaus P. Oscillating PEP. In: International physiotherapy Group for Cystic Fibrosis, editor(s). physiotherapy for people with cystic fibrosis: from infant to adults. 4th Edition. Karup: European Cystic Fibrosis Society, 2009:18–22.
39. Dasgupta B, Brown N, King M. Effects of sputum oscillations and rhDNase in vitro: A combined approach to treat cystic fibrosis lung disease. Pediatr Pulmonol 1998;26:250–55.
40. Poncin W, Reychler G, Liistro M, Liistro G. Comparison of 6 oscillatory positive expiratory pressure devices during active expiratory flow. Respi Care 2020;65:492–99.
41. Rutland J, Griffin W, Cole P. Human ciliary beat frequency in epithelium from intrathoracic and extrathoracic airways. Am Rev Respir Dis 1982;125:100–05.
42. King M, Phillips D, Gross D, Vartian V, Chang H, Zidulka A. Enhanced tracheal mucus clearance with high frequency chest wall compression. Am Rev Respir Dis 1983;128:511–15.
43. King M. Particle deposition and mucociliary clearance: Physical signs. In: Bates DV, editor. Respiratory function in disease. Philadelphia: WB Saunders; 1988:67–80.
44. Dos Santos AP, Guimaraes RC, De Carvalho EM, Gastaldi AC. Mechanical behaviours of Flutter, VRP1, Shaker and Acapella devices. Respir Care 2013;58:298–304.
45. Allam N, Khalaf M, Thabet W, Ibrahim Z. Effect of combination of acapella device and breathing exercises on treatment of pulmonary complications after upper abdominal surgeries. J Surg Special Issue Gastrointestinal Surg Recent Trends 2016;4:10–14.
46. Lu X, Qiang Y. The effect of Acapella trainer on respiratory function of patients after thoracoscopic lung cancer surgery. Precision Med Sci 2021;10:86–89.
47. Chaudhary S, Ingale Chaudhary N, Ghewade B, Mahajan G. The immediate effects of breathing exercises with Acapella and incentive spirometer on preventing early pulmonary complications following CABG — A comparative study. Int J Curr Res Rev 2020;12:51–58.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 42, No. 2 (2022) 91–97
DOI: 10.1142/S1013702522500093

Hong Kong Physiotherapy Journal
https://www.worldscientific.com/worldscinet/hkpj

Chronic knee osteoarthritis: Relationships of body mass index and selected psychosocial factors among Nigerians

Adesola Odole[1,*], Ezinne Ekediegwu[1,2,†] and E. N. D. Ekechukwu[3,‡]

[1]*Physiotherapy Department, Faculty of Clinical Sciences*
College of Medicine, University of Ibadan, Ibadan, Nigeria

[2]*Department of Medical Rehabilitation (Physiotherapy)*
Faculty of Health Sciences and Technology
Nnewi Campus, Anambra, Nigeria

[3]*Department of Medical Rehabilitation*
Faculty of Health Sciences and Technology
College of Medicine, University of Nigeria, Enugu, Nigeria

*adesola_odole@yahoo.com; acodole@com.ui.edu.ng
†tshantec@yahoo.com; ec.ekediegwu@unizik.edu.ng
‡nelson.ekechukwu@unn.edu.ng

Received 26 May 2020; Accepted 19 April 2022; Published 22 June 2022

Background: Knee Osteoarthritis is the most commonly affected joint among Africans. There is a shred of preliminary evidence that a high body mass index (BMI) is associated with high kinesiophobia. Little is known about the relationships of psychosocial factors such as Kinesiophobia, Pain Catastrophizing (PC), Self-Efficacy (SE), and BMI among Nigerians with knee OA.
Objective: This study aims to determine the relationships between BMI and selected psychosocial factors (kinesiophobia, pain catastrophizing, and self-efficacy) among individuals with knee OA in Nigeria.
Methods: Seventy-seven consecutively sampled patients diagnosed with knee OA from three selected public hospitals in Enugu, South-East Nigeria, participated in this cross-sectional survey. Brief Fear of Movement Scale for Osteoarthritis (BFMSO), Pain Catastrophizing Scale (PCS), and Arthritis Self-Efficacy Scale-8 item

†Corresponding author.

Copyright@2022, Hong Kong Physiotherapy Association. This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) License which permits use, distribution and reproduction, provided that the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

(ASES-8 item) were used to assess Kinesiophobia, PC, and SE, respectively. Also, a stadiometer and weighing scale were used to determine height and weight respectively. Data were analyzed using Pearson's correlation coefficient at $p < 0.05$ and multiple linear regression.
Results: Participants were aged 58.04 ± 12.46 years. Female participants had a higher BMI (31.51 ± 6.82) than the males (26.86 ± 3.03). The mean scores for BMI of the right knee, left knee, and bilateral knees were 29.00 ± 5.35, 24.78 ± 3.74, and 33.02 ± 6.80, respectively. Significant positive correlations were found between BMI and PC ($r = 0.35$) whereas significant negative correlations existed between BMI and SE ($r = -0.30$). Significant predictive markers of BMI were PC ($\beta = 0.21$) and SE ($\beta = -0.89$).
Conclusion: Body mass index, PC, and SE correlate significantly in individuals with knee OA. The results call for the routine integration of psychologically-informed physiotherapy practice in the management of knee OA.

Keywords: Body mass index; kinesiophobia; pain catastrophizing.

Introduction

Osteoarthritis (OA) of the knee causes the greatest burden to the adult population through disability which is the resultant effect of joint pain and stiffness.[1] Accordingly, it is the most common form of OA among Africans.[1] In Nigeria, it accounts for 65–78% of cases in hospitals.[2–4] Epidemiological studies have revealed that there are both endogenous or systemic (such as age, gender and genetic) and exogenous or local (for example obesity, microtrauma, knee joint alignment, repetitive use of joints, bone density, muscle weakness, and joint laxity play) risk factors for knee OA.[5–7] One of the strongest and best-established modifiable exogenous potent risk factors is also included but is not limited to overweight and obesity.[5,8,9]

The literature suggested that the relationship between obesity {Body Mass Index (BMI) > 30} and knee OA is stronger than with hip OA.[7] The Chingford study showed that for every two-unit increase in BMI (approximately 5 kg), the odds ratio for developing radiographic knee OA increased by 1.36.[10,35] A strong association exists between high BMI and the incidence of knee OA. Higher BMI values (greater or equal to 30 kg/m^2) negatively impact the musculoskeletal system as mechanical stress and inflammation increase in joints and body tissues result in pain and limitation to physical activities.[11–13] Furthermore, these BMI values are associated with greater adverse effects such as a higher risk of several negative psychological alterations, especially, among obese women; probably due to increased societal pressures on women to be thin.[14] These psychological discomforts include but are not limited to stigma, fear of falling, low self-efficacy pain catastrophizing (PC) and fear of movement particularly due to joint pain.[13,15–17]

Preliminary evidence reveals that high BMI (morbid obesity) is associated with high kinesiophobia.[13,34] More so, higher PC relates to lower SE for pain control, physical function emotional symptoms as mediators among individuals with high BMI values (between 25 kg/m^2 and 42 kg/m^2).[18] It has been reported that Nigerians, generally have low physical activity levels.[19–21] We, therefore, hypothesized that owing to knee OA and the resultant avoidance of physical activity, disability, and worsening of pain,[13] there would be lower levels of psychological health, especially among obese individuals with knee OA in Nigeria. Published studies on the relationships between BMI and kinesiophobia, pain catastrophizing, SE among individuals with knee OA are limited particularly in Nigeria. This study was, therefore, designed to investigate the relationships between BMI and selected psychosocial factors (kinesiophobia, pain catastrophizing, and self-efficacy) among patients diagnosed with OA of the knee in Nigeria.

Method

Ethical approval was sought and obtained from the University of Ibadan/University College Hospital Health Research Ethics Committee as well as the selected hospitals (University of Nigeria Teaching Hospital, Parklane and National Orthopaedic hospitals) for data collection before the commencement of the study.

Participants

Consecutive sampling technique was used to recruit 77 patients who fulfilled all the inclusion criteria. After recruitment, all patients consented to participate in the study and therefore, completed demographic/clinical data forms and the selected instruments. They were Igbo-literate or English-literate patients with chronic clinical and radiological features of only knee OA using the American College of Rheumatology (ACR) Clinical Classification Criteria for OA of the knee. Patients with prior knee surgery, acute knee trauma, and any other form of arthritis or intra-articular corticosteroid injection to the knee(s) three weeks before recruitment for the study were excluded from the study. Each participant's height (in meters) and weight (in kilogram) were measured to determine their BMI [m/kg^2].

Measures

- Kinesiophobia

The **Brief Fear of Movement Scale for Osteoarthritis (BFMSO)** is an adapted version of the Tampa Scale of Kinesiophobia (TSK) for patients with OA which consists of six items. The scale has sound psychometric properties including convergent validity.[22] Possible scores range from 6 to 24, with higher scores ($\geq$15) representing a high degree of kinesiophobia.[23]

- Pain Catastrophizing

The **Pain Catastrophizing Scale (PCS)** was used as a measure of pain catastrophizing. The PCS is a 13-item self-report measure with three subscales of magnification, rumination, and helplessness.[24] The PCS has excellent psychometric properties, including adequate to excellent internal consistency,[25–27] test–retest reliability, good convergent validity, and constructs validity. Total score ranges from 0 to 52 with scores 30 and above representing a clinically relevant level of catastrophizing.[25]

- Self-efficacy

SE was assessed with **Arthritis Self-efficacy Scale-8 Item (ASES-8)**, the shortened form of the original ASES 20-item which consists of eight items with no subscales. The total score is the mean of the eight items with higher scores denoting greater SE.[28] The scale demonstrates high internal consistency[28] and positively correlates with other measures of self-rated health status and physical performance but negatively correlates with arthritis symptoms (pain, fatigue, and stiffness).[28]

- Anthropometric characteristics

Each participant's height was measured in centimeters (converted to meters) and weight was measured without shoes to the nearest 1 kg using a weighing scale balance (electronic digital) with a height meter (Zt-120). BMI was calculated from these measures using the formula: BMI (kg/m^2) = weight (kg)/(height (m))^2.

- Demographic characteristics

Age, gender and joint affected were assessed using a demographic questionnaire.

Data analysis

Categorical variables (gender and joint affected) were represented using counts and proportions (%). Descriptive statistics (mean and SD) were calculated for selected psychosocial variables (kinesiophobia, pain catastrophizing, and self-efficacy) and BMI for each category. Pearson's correlation coefficient was used to determine the correlation between the scores of each selected psychosocial variable and BMI. The closer the coefficient is to zero (from either + or −), the less strong the relationship and vice versa.[29] From this study, the correlation coefficient (r) values for most of the variables were weak (very close to zero: 0.20–0.39). The relationships would have been stronger if the correlation coefficient were closer to one. Multiple linear regression analysis was used to determine the effect of each of the independent/predictor variables (selected psychosocial factors) on the dependent (outcome) variable (BMI). The level of significance was set at 0.05.

Results

Summary of participant's characteristics

Seventy-seven (15 males, 62 females) consecutively sampled patients with knee OA (mean age: 58.04 $\pm$ 12.46 years) participated in this study. The mean scores of Kinesiophobia, Pain Catastrophizing, and SE of all the participants were 14.05 $\pm$ 3.61, 16.43 $\pm$ 9.99, and 7.53 $\pm$ 1.87, respectively. Overall

Table 1. Characteristics of participants and mean scores of body mass index, kinesiophobia, pain catastrophizing, and SE across descriptive variables.

Variables	Categories	Mean scores					
		Frequency	Percentage (%)	BMI (kgm^2)	K	PC	SE
Sex	Male	15	19.5	26.86 ± 3.03	12.53 ± 3.58	13.13 ± 12.73	7.40 ± 1.90
	Female	62	80.5	31.51 ± 6.82	14.42 ± 3.55	17.23 ± 9.16	7.56 ± 1.88
Joint affected	Right only	32	41.6	29.00 ± 5.35	13.19 ± 3.72	18.34 ± 10.35	7.87 ± 1.72
	Left only	7	9.1	24.78 ± 3.74	15.29 ± 3.35	8.00 ± 5.54	8.36 ± 1.01
	Right and left	38	49.4	33.02 ± 6.80	14.55 ± 3.49	16.37 ± 9.68	7.09 ± 2.03
Total		77	100	30.60 ± 6.51	14.05 ± 3.61	16.43 ± 9.99	7.53 ± 1.87

Table 2. Relationship among body mass index, kinesiophobia, pain catastrophizing, and self-efficacy: Pearson's correlation matrix.

Variables		Correlation coefficients			
		Body mass index	Kinesiophobia	Pain catastrophizing	Self-efficacy
Body mass index	**r**	1	0.06	0.35*	−0.30*
	p	0.60	0.00	0.01	
Kinesiophobia	**r**	0.06	1	0.28*	−0.01
	p	0.60	0.01	0.92	
Pain catastrophizing	**r**	0.35*	0.28*	1	−0.28*
	p	0.00	0.01		0.01
Self-efficacy	**r**	−0.30*	−0.01	−0.28*	1
	p	0.01	0.92	0.01	

Notes: r - Correlation coefficient, p - Probability of error, and * - Significant Correlation at $p < 0.05$.

Table 3. Variables associated with BMI in the multivariate analyses.

Outcome variables	Predictor variables	β	Standard error	t	p	F	R^2
Body mass index							
	Kinesiophobia	−0.08	−0.05	−0.42	0.68	5.69 ($p < 0.05$)	0.19
	Pain catastrophizing	0.21*	0.33	2.96	0.00		
	Self-efficacy	−0.89*	0.37	−2.40	0.02		

Notes: β = Contribution of each variable (Unstandardized coefficient for applied studies), R^2 = Quality of the fitness of the models; coefficient of determination, F = Joint significance of all the variables in the model, p = Probability of error, and * = Significant predictor variables at $p < 0.05$.

average BMI was $30.60 \pm 6.51\,\text{kg/m}^2$. For joint affectation, the mean BMI of the right knee only, left knee and bilateral knee were $29.00 \pm 5.35\,\text{kg/m}^2$, $24.78 \pm 3.74\,\text{kg/m}^2$, $33.02 \pm 6.80\,\text{kg/m}^2$, respectively as shown in Table 1. The BMI of female participants ($31.51 \pm 6.82\,\text{kg/m}^2$) was more than the male ($26.86 \pm 3.03\,\text{kg/m}^2$). Furthermore, women reported higher scores on kinesiophobia (14.42 ± 3.55), PC (17.23 ± 9.16), and SE (7.56 ± 1.88) than the males (12.53 ± 3.58, 13.13 ± 12.73, 7.40 ± 1.90, respectively). Significant positive correlations were found between BMI and PC($r = 0.35$) whereas significant negative correlations existed between BMI and SE($r = -0.30$) (Table 2). Significant predictive markers of BMI were PC($\beta = 0.21$) and SE($\beta = -0.89$) (Table 3).

Discussions

This study appears to be the first to examine the relationship between BMI and selected psychosocial factors among individuals with knee OA in Nigeria. The mean age (58.04 ± 12.46 years) of the participants from this study supports the definition of OA by the American College of Rheumatology as a disease that most often affects middle-aged to elderly people.[30] One of the most interesting findings of this study was that the BMI of female participants appeared higher (class I obesity) than that of the male participants (overweight), even though both genders had BMI values above the normal range. On a general note, the average BMI of all the participants was high (30.60 ± 6.51); this reveals that Nigerians with knee OA (particularly middle-aged) appear to be obese. More so, the mean scores of the selected psychosocial factors were higher among women than men. This is not consistent with previous studies whereby no difference in PC was found between men and women with knee OA and on kinesiophobia, men reported higher scores than women.[31,32] The mean scores of both kinesiophobia and PC (14.05 ± 3.61, 16.43 ± 9.99, respectively) of the participants appeared clinically insignificant. However, the mean BMI score for bilateral knee OA was observed to be higher (mildly obese) than unilateral OA. A significant positive correlation was found between BMI and PC ($r = 0.35$, $p \leq 0.01$) whereas a significant negative correlation existed between BMI and SE ($r = -0.30$, $p \leq 0.01$). This supports the findings from the study of Somers and colleagues[33] that both BMI and PC are directly proportional. An insignificant correlation was found between kinesiophobia and BMI ($r = 0.06$, $p \leq 0.60$). This suggests that the high mean BMI value of the participants was not significant enough to lead to kinesiophobia. Taking together, these findings suggest that BMI though relates statistically significant with pain catastrophizing, but in a clinically insignificant manner. Also, high BMI in knee OA has a negative influence on self-efficacy.

Our findings regarding the relationship between BMI and kinesiophobia, pain catastrophizing, SE among patients with knee OA have two important clinical implications. First, they suggest that clinicians working with patients with knee OA should be more aware of their BMI. It could be useful, for example, to include an assessment of kinesiophobia, pain catastrophizing, and SE in the clinical evaluation of patients with high BMI. Secondly, identifying patients with increased BMI who are prone to these psychosocial factors could lead to timelier referral of patients for interventions designed to reduce both BMI and psychosocial factors. BMI is a modifiable risk factor of knee OA and reductions in this could potentially decrease PC and enhance the SE of individuals with knee OA. It would be pertinent to study the relationship between kinesiophobia and BMI in classes II and III obese individuals with knee OA, as no significant relationship existed between the two variables in this study, that is, among overweight and class I obese individuals. Studies on the relationship between BMI and other psychosocial factors are needful.

Summary

This study appears to be the first to demonstrate significant relationships amongst BMI, PC and SE in Nigerians with knee OA. BMI has psychosocial effects, especially, among individuals with knee OA in Nigeria. The predominance of knee OA and increased BMI value among Nigerian women with knee OA calls for a more detailed and adequate assessment and management. Our findings regarding the relationships between BMI and pain catastrophizing, and SE raises the intriguing possibility that robust assessment of BMI and these psychosocial factors in the management of knee OA could lead to better diagnosis with resultant optimal care. In addition to this, Psychologically-Informed Physiotherapy Practice (PIPP) would be ideal in the management of patients with knee OA who have an overlay of these psychosocial factors.

Acknowledgments

The authors thank all patients with knee OA in the selected public hospitals in Enugu, South-East Nigeria for their kind participation in this study. The authors sincerely appreciate the research assistants (physiotherapists), orthopaedic surgeons, and nurses for their assistance in the recruitment recruitment process.

Ethical Approval

University of Ibadan/University College Hospital Health Research Ethics Committee (UI/EC/15/0059).

Conflict of Interest

None declared.

Funding/Support

The study did not receive external funding.

Author Contributions

AO conceptualized the idea, supervised the study, and reviewed the manuscript.

EE collected and coded the data, took part in data analysis, and prepared the manuscript. EE was also involved in the conceptualisation.

END analyzed the data and reviewed the manuscript.

References

1. Anna L, Mark E, Elaine D, Cyrus C. Epidemiology, Burden of Osteoarthritis. Br Med Bull 2013;105:185–99.
2. Ebong WW. Osteoarthritis of the knee in Nigeria. Ann Rheum Dis 1985;44:682–84.
3. Ogunlade SO, Alonge TO, Omololu AB, Adekolujo OS. Clinical spectrum of large joint osteoarthritis in Ibadan, Nigeria. Eur J Sci Res 2005;11:116–22.
4. Akinpelu AO, Alonge OO, Adekanla BA, Odole AC. Patterns of osteoarthritis seen in physiotherapy facilities in Ibadan, Lagos, Nigeria. Afr. J. Biomed. Res. 2007;10:111–15.
5. Moseley JB, O'Malley K, Petersen NJ. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. New Engl J Med 2002;347:81–88.
6. Reid CR, Bush PM, Cummings NH, McMullin DL, Durrani SL. A review of occupational knee disorders. J Occup Rehabil 2010;20:489–01.
7. Kumar P and Clark M. Kumar and Clark's clinical medicine. Spain: Elsevier Ltd, 2012.
8. Christensen R, Bartels EM, Astrup A, Bliddal H. Effect of weight reduction in obese patients diagnosed with knee osteoarthritis: A systematic review and meta-analysis. Ann Rheumatic Dis 2007;66 (4):433–9.
9. Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Obesity and osteoarthritis in knee, hip and/or hand: An epidemiological study in the general population with 10 years follow-up. BMC Musculoskeletal Disord 2008;9(1):132.a.
10. Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Prevalence and burden of osteoarthritis: Results from a population survey in Norway. J Rheumatol 2008;35(4):677–84.
11. Cesari M, Kritchevsky SB, Baumgartner RN, Atkinson HH, Penninx BW, Lenchik L, Palla SL, Ambrosius WT, Tracy RP, Pahor M. Sarcopenia, obesity, and inflammation — results from the trial of angiotensin-converting enzyme inhibition and novel cardiovascular risk factors study–. Amer J Clin Nutr 2005;82(2):428–34.
12. Vincent KR, Vincent HK. Resistance exercise for knee osteoarthritis. PM R 2012;4(5 Suppl): S45–S52.
13. Vincent HK, Adams MCB, Vincent KR, Hurley RW. Musculoskeletal pain, fear-avoidance behaviors, and functional decline in obesity: Potential interventions to manage pain and maintain function. Reg Anesth Pain Med 2013;38(6):481–91.
14. Carpenter KM, Hasin DS, Allison DB, Faith MS. Relationships between obesity and DSM-IV major depressive disorder, suicide ideation, and suicide attempts: Results from a general population study. Am J Public Health 2000;90(2):251–57.
15. Shelby RA, Somers TJ, Keefe FJ, Pells JJ, Dixon KE, Blumenthal JA. Domain specific self-efficacy mediates the impact of pain catastrophizing on pain and disability in overweight and obese osteoarthritis patients. J Pain 2008;9(10):912–19.
16. Sallinen J, Leinonen R, Hirvensalo M, Lyyra T-M, Heikkinen E, Rantanen T. Perceived constraints on physical exercise among obese and non-obese older people. Prev Med 2009;49(6):506–10.
17. Groven KS, Engelsrud G. Dilemmas in the process of weight reduction: Exploring how women experience training as a means of losing weight. Int J Qual Stud Health Well-Being 2010;22:5.
18. Shelby RA, Somers TJ, Keefe FJ, Pells JJ, Dixon KE, Blumenthal JA. Domain specific self-efficacy mediates the impact of pain catastrophizing on pain and disability in overweight and obese osteoarthritis patients. J Pain 2008;9(10):912–19.
19. Odunaiya NA, Ayodele OA, Oguntibeju OO. Physical activity levels of senior secondary school students in Ibadan, western Nigeria. West Ind Med J 2010;59(5):529–34.
20. Adeniyi AF, Okafor NC, Adeniyi CY. Depression and physical activity in a sample of Nigerian adolescents: Levels, relationships and predictors. Child Adolesc Psychiatry Ment Health 2011;5(1):16.
21. Oyewole OO, Odusan O, Oritogun KS, Idowu AO. Physical activity among type-2 diabetic adult Nigerians. Ann African Med 2014;13(4):189.

22. Shelby, RA, Somers, TJ, Keefe, FJ, DeVellis, BM, Patterson, C, Renner, JB and Jordan, JM. Brief fear of movement scale for osteoarthritis. Arthritis Care Res, 2012;64:862–871. https://doi.org/10.1002/acr.21626.
23. Shelby Rebecca A, Somers Tamara J, Keefe Francis J, DeVellis Brenda M, Patterson Carol, Renner Jordan B, Jordan Joanne M. Brief fear of movement scale for osteoarthritis. Arthritis Care Res (Hoboken). 2012;64(6):862–71.
24. Sullivan MJ, Bishop SR, Pivik J. Pain Catastrophizing scale. The Journal of Pain. 2009.
25. Bansal D, Gudala K, Lavudiya S, Ghai B, Arora P. Translation, adaptation, and validation of Hindi version of the pain catastrophizing scale in patients with chronic low back pain for use in India. Pain Medicine. 2016 Oct 1;17(10):1848–58.
26. Nishigami T, Mibu A, Tanaka K, Yamashita Y, Watanabe A, Tanabe A. Psychometric properties of the Japanese version of short forms of the Pain Catastrophizing Scale in participants with musculoskeletal pain: A cross-sectional study. Journal of Orthopaedic Science. 2017 Mar 1;22(2):351–6.
27. Wheeler CH, Williams AC, Morley SJ. Meta-analysis of the psychometric properties of the Pain Catastrophizing Scale and associations with participant characteristics. Pain. 2019 Sep 1;160 (9):1946–53.
28. Brady, TJ. Measures of self-efficacy: Arthritis Self-Efficacy Scale (ASES), Arthritis Self-Efficacy Scale-8 Item (ASES-8), Children's Arthritis Self-Efficacy Scale (CASE), Chronic Disease Self-Efficacy Scale (CDSES), Parent's Arthritis Self-Efficacy Scale (PASE), and Rheumatoid Arthritis Self-Efficacy Scale (RASE). Arthritis Care & Research, 2011;63(S11):S473–S485.
29. Pallant J. SPSS survival manual: A step by step guide to data analysis using IBM SPSS. Routledge; 2020 Jul 16.
30. Hochberg MC, Altman RD, April KT, Benkhalti M, Guyatt G, McGowan J, Towheed T, Welch V, Wells G, Tugwell P. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis care & research. 2012 Apr;64(4):465–74.
32. Bränström H, Fahlström M. Kinesiophobia in patients with chronic musculoskeletal pain: differences between men and women. Journal of rehabilitation medicine. 2008 May 1;40(5):375–80.
31. Tonelli SM, Rakel BA, Cooper NA, Angstom WL, Sluka KA. Women with knee osteoarthritis have more pain and poorer function than men, but similar physical activity prior to total knee replacement. Biology of sex differences. 2011 Dec;2 (1):1–2.
33. Somers TJ, Keefe FJ, Carson JW, Pells JJ, LaCaille L. Pain catastrophizing in borderline morbidly obese and morbidly obese individuals with osteoarthritic knee pain. Pain Res Manag 2008;13(5):401–6.
34. Vincent HK, Lamb KM, Day TI, Tillman SM, Vincent KR, George SZ. Morbid obesity is associated with fear of movement and lower quality of life in patients with knee pain-related diagnoses. PM&R 2010;2(8):713–22.
35. Hart DJ, Spector TD. The relationship of obesity, fat distribution and osteoarthritis in women in the general population: The Chingford Study. J Rheumatol 1993;20(2):331–35.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 42, No. 2 (2022) 99–110
DOI: 10.1142/S101370252250010X

Hong Kong Physiotherapy Journal
https://www.worldscientific.com/worldscinet/hkpj

Use of mobile app to enhance functional outcomes and adherence of home-based rehabilitation program for elderly with hip fracture: A randomized controlled trial

Kui Ching Cheng[1,*], Kin Ming Ken Lau[2], Andy S. K. Cheng[3], Tin Sing Keith Lau[1], Fuk On Titanic Lau[1], Mun Cheung Herman Lau[4] and Sheung Wai Law[5]

[1]*Physiotherapy Department, Tai Po Hospital, 9 Chuen on Road, Tai Po New Territories, Hong Kong SAR, China*

[2]*Physiotherapy Department, Tung Wah Eastern Hospital 19 Eastern Hospital Road, Causeway Bay, Hong Kong SAR, China*

[3]*Department of Rehabilitation Sciences The Hong Kong Polytechnic University, 11 Yuk Choi Road Hung Hom, Kowloon, Hong Kong SAR, China*

[4]*CUHK Medical Centre, 9 Chak Cheung Street Shatin, New Territories, Hong Kong SAR, China*

[5]*Department of Orthopaedics and Traumatology CUHK Medical Centre, 9 Chak Cheung Street Shatin, New Territories, Hong Kong SAR, China*

*ckc863@ha.org.hk

Received 9 December 2021; Accepted 11 April 2022; Published 15 June 2022

Background: Mobile app has been used to improve exercise adherence and outcomes in populations with different health conditions. However, the effectiveness of mobile app in delivering home-based rehabilitation program to elderly patients with hip fracture is unclear.

*Corresponding author.

Copyright@2022, Hong Kong Physiotherapy Association. This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) License which permits use, distribution and reproduction, provided that the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Objective: The aim of this study was to test the effectiveness of mobile app in delivering home-based rehabilitation program for improving functional outcomes and reducing caregiver stress with enhancing adherence among the elderly patients with hip fracture.
Methods: A randomized controlled trial with an intervention period of two months was performed. Eligible participants were randomized into either experimental group with home-based rehabilitation program using a mobile app or control group with home-based rehabilitation program using an exercise pamphlet. Primary outcomes were Modified Functional Ambulatory Category (MFAC), Elderly Mobility Scale (EMS) and Lower Extremity Functional Scale (LEFS). Secondary outcomes were exercise adherence and Modified Caregiver Strain Index (M-CSI). The outcomes were collected at pre-discharge training session, one month and two months after hospital discharge.
Results: A total of 50 participants were enrolled, with 19 participants in the experimental group and 20 participants in the control group. Eleven participants had withdrawn from the study. The experimental group showed higher exercise adherence than the control group in first month ($p = 0.03$). There were no between-group differences in MFAC, EMS, LEFS and M-CSI at the first month and second month.
Conclusion: Use of the mobile app improved exercise adherence, yet it did not improve physical performance, self-efficacy and reduce caregiver stress when compared to a standard home rehabilitation program for elderly patients with hip fracture. Further studies to investigate the benefits of mobile apps are required. (ClinicalTrials.gov ID: NCT04053348.)

Keywords: Caregivers; exercise; hip fractures; mobile application.

Introduction

About half of women and a quarter of men suffer from a fragility bone fracture in their lifetime,[1,2] which is mostly associated with fall.[2] Hip fracture is commonly seen among all fragility fracture cases and it places an increasing burden on healthcare service providers around the world. A global estimation on the number of hip fractures in 2050 is between 7.3 million and 21.3 million,[3] and over half of them will occur in Asia due to the higher growth of ageing population in many Asian regions.[4] A 2015 study showed that there was a steady increase in the incidence of geriatric hip fracture in Hong Kong.[5] Assuming no increase in age-specific rates, the estimated annual incidence of elderly hip fracture in 2040 is more than 14,500.[5]

Previous studies showed that post-fracture limitations, such as impaired balance,[6] decreased mobility and lower limb strength,[6,7] restrained elderly patients with hip fracture from participating in daily activities to live independently and safely.[6,8] Functional performance of these patients was greatly affected by these limitations. More than half of them failed to regain pre-fracture functional abilities two years after injury and they have a moderate-to-high risk of further falls.[8,9] Main factors affecting functional performance were pre-fracture ambulatory level and age.[10] Post-fracture limitations also increase the burden of informal caregivers, who are responsible for providing assistance to patients in performing their daily activities safely. A study reported that 50% and 26% of caregivers perceived a high level of caring stress at one month and one year after the patients were discharged from hospital, respectively.[11] Factors influencing caregiver stress level, such as pre-fracture functional level, caregiver–patient relationship and age of patients, were also identified in the study.[11]

Provision of educational materials, which consist of home exercises and caring skills, is essential for patients and caregivers to take responsibility and manage their conditions at home. Adherence to home exercises has been shown to be associated with improved patient outcomes.[12,13] Educational materials are conventionally delivered through paper handouts, which are reliable and convenient for distribution. However, written texts and illustrations could be incomprehensible, leading to a decrease in patients' exercise adherence at home.[14] Other possible factors, such as self-efficacy, age and perceived social support from family, were shown to be associated with exercise adherence.[15]

Use of mobile apps in exercise education was shown to improve exercise knowledge and self-efficacy when compared with the use of paper handouts.[16] Some researchers have therefore directed their attention in evaluating the strategy

of using technology to improve patients' exercise adherence and outcomes. Recent randomized controlled trials (RCTs) reported greater adherence to home programs, which were delivered with an Internet-based self-monitoring system via mobile phones, in patients with hemophilia-related knee dysfunction and musculoskeletal conditions, respectively.[17,18] Promising results were also reported with the use of mobile apps to improve adherence and outcomes in other health areas, such as weight loss and diabetic management.[19,20] However, elderly patients may find difficulties in using mobile apps and experience usability issues, such as limited screen size of mobile devices and complicating app interactions.[21] Successful application of mobile apps in populations with different health conditions may not be transferable to geriatric populations with hip fracture. Thus, a joint research team with the collaboration among Physiotherapy Department of Tai Po Hospital (TPH) in New Territories East Cluster (NTEC), Hong Kong Polytechnic University Rehabilitation Science (HKPURS) Department and NTEC Information Technology Department was formed to conduct a study to test the effectiveness of mobile app in delivering home-based rehabilitation program for improving functional outcomes and reducing caregiver stress with enhancing adherence among the elderly with hip fractures.

The study's primary hypothesis is that the home-based rehabilitation program delivered using mobile app will result in better functional outcomes compared to usual care delivered using the conventional paper handouts. Secondary hypothesis is that mobile app intervention will result in better adherence and lesser caregiver burden.

Method

Study design

It was a parallel, two-arm (experimental control), RCT with a two-month intervention period (see Fig.1).

Participants

All patients were recruited from the inpatient Geriatric Hip Fracture Rehabilitation Program in the Department of Orthopaedic Rehabilitation (DOR) of Tai Po Hospital. Tai Po Hospital is a tertiary-care hospital and it provides rehabilitation services to hip fracture patients who are transferred from acute hospitals in the New Territories East Cluster. To be eligible for enrolment, patients fulfilled the following inclusion criteria: (1) have a primary diagnosis of hip fracture; (2) be 60–90 years old; (3) be literate enough to read and understand simple questions in Chinese; (4) be discharged home and taken care by caregiver; (5) have at least category three measured by the Modified Functional Ambulatory Category (MFAC) upon discharge; (6) have access to a smartphone or tablet (iOS or Android platform); (7) have signed the written informed consent. The exclusion criteria were as follows: (1) have a bilateral hip fracture or hip fracture is the result of a malignancy; (2) have significant cognitive deficits with a score of less than 19 in Hong Kong Version of Montreal Cognitive Assessment (HK-MoCA)[22]; (3) have significant cardiopulmonary contraindications or pre-existing conditions that preclude participation in an exercise program; (4) have a terminal illness (survival expected to be less than one year); (5) have severe visual deficits or legally blind.

Based on the power being set at 0.80, type-1 error α of 0.05 and previous similar studies,[23,24] the sample size needed in this design to run two-way repeated measures analysis of variance (ANOVA) was 126 (63 participants per group). This sample achieved 80% power to test the main effect if Geisser–Greenhouse Corrected F-test was used with a 5% significance level and an effect size of 0.42. It also achieved 82% power to test the interaction effect if Geisser–Greenhouse Corrected F-test was used with a 5% significance level and an effect size of 0.41. With the consideration of possible drop-out of the participants (an estimated 15% attrition), a total of 144 participants (72 participants in each group) should be recruited for the whole study.

Potential participants were identified by case physiotherapists a week before discharge and these participants received brief information about the study. Research team members approached the participants who agreed to be contacted for further explanation of the study. A written informed consent was obtained from each participant before enrolment. Once enrolled, case physiotherapists assessed the participants to collect baseline characteristics. The participants were then allocated to either experimental group or control group using a computer-generated randomization sequence. Ethical approval was obtained from the Joint

Chinese University of Hong Kong–New Territories East Cluster Clinical Research Ethics Committee before start of the study.

Potential risk and preventive measures

All participants recruited were screened by physiotherapists according to the exercise prescription guidelines of the American College of Sports Medicine (ACSM) for any contraindications.[25] Suitability of participating in home exercise programs was assessed by physiotherapists based on environmental risk, fall risk and competence of participants or caregivers in performing home exercises. Participants should have sufficient space and appropriate supporting furniture at home for performing the exercises. Both participants and caregivers should also have a good understanding of prescribed exercises so that participants could perform the exercises safely. Caregivers should be able to supervise patients in performing home exercises.

Intervention

A briefing session was arranged for all participants with their caregivers before hospital discharge. This session included provision of home-based rehabilitation program and caregiver education by case physiotherapists. For participants and caregivers in the experimental group, they were instructed to install the mobile app in the participants' mobile devices. A smart tablet (240 mm × 169.5 mm), which also had the mobile app installed, could be loaned to those participants without any suitable mobile devices. They learnt how to operate the mobile app and perform exercises along with exercise videos embedded in the app. For participants and caregivers in the control group, they received an exercise pamphlet and a caregiver skill pamphlet instead of the mobile app. The pamphlets consisted of written texts and printed illustrations. Both groups received the same home-based rehabilitation program and caregiver skill information. The caregiver education involved demonstration on assisting patient in transfer and ambulation (Table 1). Physiotherapists assessed the participants during the briefing session and if there were any tasks that required assistance to complete safely, physiotherapists would instruct the caregivers to view the corresponding educational material in the mobile app or pamphlet.

Table 1. Tasks included in the caregiver education.

Category	Task
A. Patient Transfer	From lying to sitting From sitting to standing Bed-to-chair transfer
B. Patient Ambulation	Level ground walking Stairs walking

The home-based rehabilitation program for hip fracture patients involved a combination of training focused on strength, coordination and functional movements of geriatric hip fracture patients. All exercises in the rehabilitation program were safe for patients who underwent different hip fracture operations. There were four levels of difficulty, consisting of four–six exercises in each exercise level (Table 2). Participants were instructed to perform the prescribed level of exercises once per day with the course length varying from 20 min to 30 min. The progression level of the exercise program was reassured by the domiciliary physiotherapists during weekly home visits based on clinical performance of the participants. To ensure safety in performing home exercises without a physiotherapist, participants were required to perform the prescribed exercises correctly in the briefing session. Literacy of the caregivers was also assessed by physiotherapists. They were required to demonstrate the use of mobile app or exercise booklet, and explain the prescribed exercises to physiotherapists at the end of the briefing session.

Mobile app development and features

A steering group including physiotherapists, informatics and university research experts was formed to co-design the app, formulate the study design and compose educational content for the patients. Meetings were also held with all involved clinicians to refine the mobile app before study commencement.

The mobile app was developed to facilitate the implementation of a home-based rehabilitation program. Participants could use this mobile app to follow home-based exercises prescribed by their physiotherapists, track their exercise progress and obtain relevant information about hip fracture rehabilitation. Their caregivers could also refer to

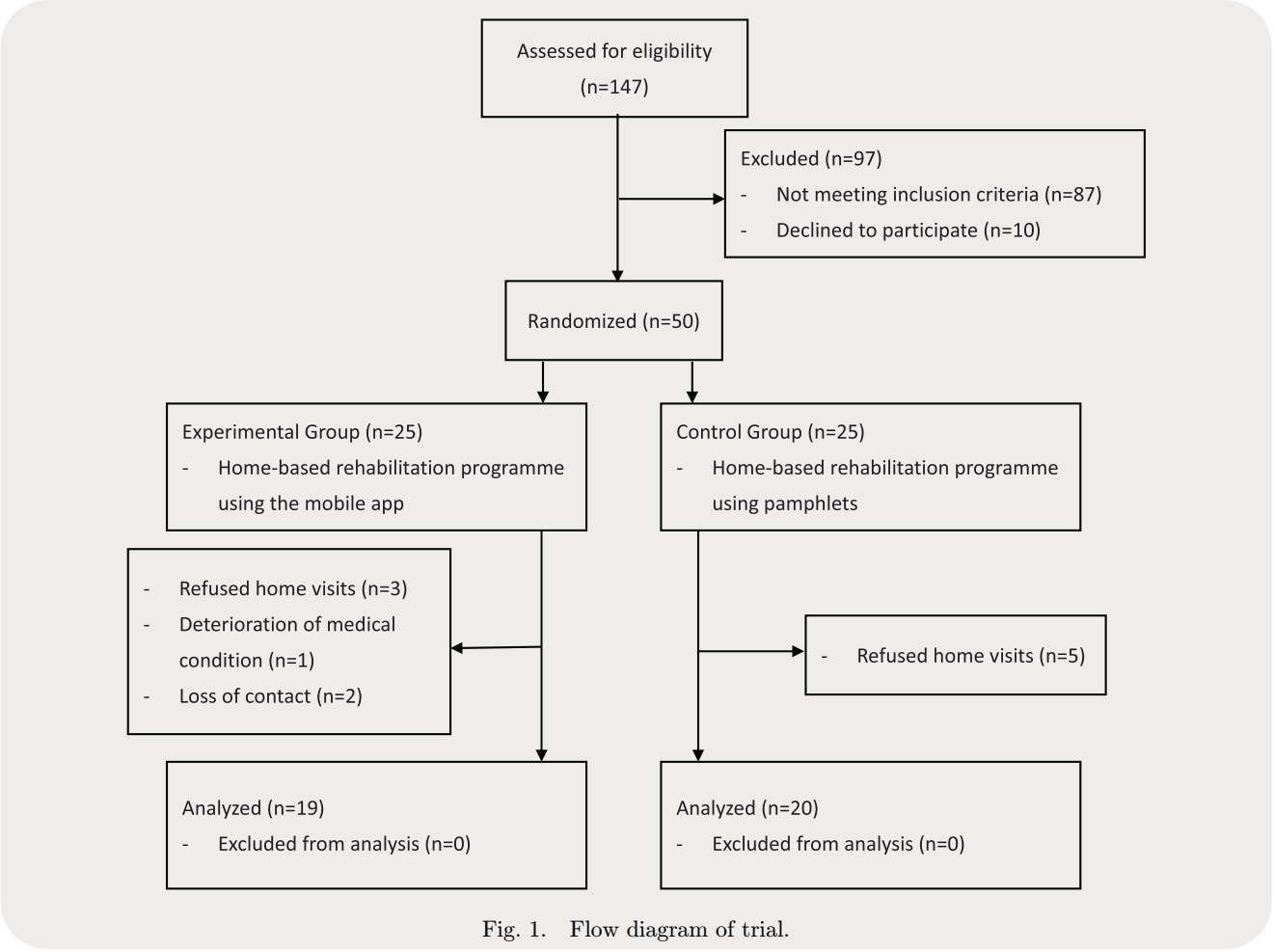


Fig. 1. Flow diagram of trial.

this information to learn about how to assist their care receivers. The features of the mobile app include the following:

(1) Exercise program: Preloaded exercises at four different levels of difficulties for the geriatric hip fracture patient (Table 2).
(2) Progress summary: It allows patients to see their progress and ultimately improve their self-efficacy through completion of tasks. A summary of the rehabilitative exercises completed by the patient will be presented in calendar format for easy interpretation.
(3) Push reminder: It alerts the participants to follow the exercise schedule and sends out motivational message to encourage them.
(4) Rehab knowledge: It allows participants and their caregivers to obtain knowledge on post-hip fracture management including the related surgical intervention, post-operative management and precautions.
(5) Caregiver skill: Video libraries of a wide range of practical skills and information about taking care of geriatric hip fracture patient for the caregivers.
(6) Support information: Useful online health information and supporting resources for participants and caregivers.

Outcomes

Demographic data of the patients including age, gender, post-operation duration and length of hospital stay was collected upon discharge. Assessments on specific outcomes were conducted at pre-discharge training session (T_0) by the case physiotherapists, first month (T_1) and second month (T_2) post-discharge by domiciliary physiotherapists during home visits. There were three primary and two secondary outcomes.

Primary outcomes

Modified functional ambulatory category

The MFAC is a seven-point Likert scale to classify a patient's walking capacity. Gait is divided into

Table 2. Four levels of difficulty in the home-based rehabilitation program.

Exercise item	Equipment needed	Training side	Dose	Levels of difficulty 1	2	3	4
1. Knee extension in sitting (resistance*)	Resistance band	Both	Ten repetitions for two sets	√	√	√	√
2. Ankle plantarflexion in sitting (resistance*)	Resistance band	Both	Ten repetitions for two sets	√	√	√	√
3. Sit-to-stand (assisted)	Table	Both	Ten repetitions for two sets	√			
4. Sit-to-stand (aided)	Table	Both	Ten repetitions for two sets		√		
5. Sit-to-stand (unaided)	Table	Both	Ten repetitions for two sets			√	
6. Knee flexion in standing	Table	Both	Ten repetitions for two sets	√			
7. Hip abduction in standing	Table	Both	Ten repetitions for two sets		√		
8. Hip abduction in standing (resistance*)	Table + Resistance band	Both	Ten repetitions for two sets			√	√
9. Mini-squat	Table	Both	Ten repetitions for two sets		√		
10. Lunge	Table	Both	Ten repetitions for two sets			√	√
11. Toe and heel raise in standing	Table	Both	Ten repetitions for two sets				√
12. Hip marching in standing	Table	Both	Thirty repetitions for two sets				√

Note: *Resistance: A resistance band was used, providing 4-lb resistance at 100% elongation.

seven categories, ranging from no ability to walk and requires manual assistance to sit or is unable to sit for 1 min without back or hand support (MFAC 1) to the ability to walk independently on level and non-level surfaces, stairs and inclines (MFAC 7).[26] The inter-rater reliability of MFAC [the intraclass coefficient (ICC)] was 0.982 (0.971–0.989), with a kappa coefficient of 0.923 and a consistency ratio of 94% for the stroke patient. The ICC of MFAC in patients with hip fractures was 0.96.[26]

Elderly mobility scale

The Elderly Mobility Scale (EMS) was used to assess an individual's mobility problems through seven functional activities including lying, sitting, standing and walking.[27] The possible total score is 20. Patients receiving a score under 10 indicate that they are dependent on mobility and activities of daily living (ADL). Patients receiving a score between 10 and 13 indicate that they are marginal in terms of safe mobility and independent in ADL. Patients receiving a score over 14 indicate that they are able to perform mobility and ADL independently and safely. EMS demonstrated good inter-rater reliability and concurrent validity.[28]

Lower extremity functional scale

Lower Extremity Functional Scale (LEFS) is a 20-item questionnaire intended to measure patients' functions with a wide range of lower extremity conditions.[29] Each item is rated on a five-point scale (0 = extreme difficulty or unable to perform activity, 4 = no difficulty); total scores range from 0 to 80, and lower scores represent greater difficulty. It has been shown to be highly reliable, correlates with other constructs and is an independent predictor of the patient and physician's assessment of change in patients.[29] The minimal clinically important difference (MCID) for the LEFS is nine points.[29] This questionnaire has been validated in Taiwan-Chinese version.[30]

Secondary outcomes

Exercise adherence

Adherence to the prescribed home-based rehabilitation program was collected in both groups at the first month (T_1) and second month (T_2) after discharge using exercise logs. Participants were asked to record their exercise logs either by the use of the mobile app for the experimental group or by exercise diaries provided for the control group. Collected exercise logs were used to calculate the percentage of the number of exercise sessions completed to the number of exercise sessions prescribed in the two-month period.

Modified caregiver strain index

The original version of the Caregiver Strain Index (CSI) was developed in 1983. It was designed to

detect the physical, psychological, social and financial strain of the informal caregiver. It consisted of 13 items. There were only two options for the respondent to choose (score 1 if the respondent chose "yes", otherwise score 0 if the respondent chose "no"). The CSI was modified later in 2003. Some of the items were rephrased and one option category (yes, sometimes) was added. There were three option categories for the respondents to choose and the score ranged from 0 to 26. Caregivers would face a high level of strain with high CSI score. The modified version of CSI [i.e. the Modified Caregiver Strain Index (M-CSI)] achieved a high internal reliability ($\alpha = 0.90$).[31] Chan *et al.* examined whether M-CSI was still valid when it was applied on the Hong Kong Chinese caregivers in 2013. They validated the Chinese version of the Modified Caregiver Strain Index and the results showed that it achieved a good internal reliability ($\alpha = 0.91$).[32]

Data analysis

Independent t-test and Pearson's Chi-square test were used to compare the baseline characteristics between the two groups. One-way repeated measures ANOVA was conducted to see the within-group differences at different assessment occasions. If there was a significant difference within a group, a post-hoc paired t-test would be conducted to evaluate this within-group difference. Additionally, a two-way repeated measures ANOVA was used to find out if there was any significant interaction effect between the two groups across different assessment occasions. If there was a significant effect between the two groups, a post-hoc independent t-test would be carried out with Bonferroni correction. All statistical analyses were performed using the IBM SPSS program version 28 for Windows; the significance level was set at $p < 0.05$.

Results

From October 2019 to March 2021, 50 participants were recruited and randomized into experimental and control groups. Eleven participants withdrew from the study, with eight participants refusing home visits, one participant having deterioration of medical condition and two participants being unable to contact. There were no adverse events that occurred as a result of participation. The proposed sample size of 144 participants was unable to be achieved in this study due to a significant decrease in the number of eligible participants in the hospital after the outbreak of COVID-19. High dropout rates of participants due to refusal of home visits were observed. Thirty-nine participants completed the study and data was collected for analysis, including 19 participants in the experimental group and 20 participants in the control group. Data normality was checked using the Shapiro–Wilk test. Since the test results suggested non-normality, non-parametric tests for data analysis were adopted. Mann–Whitney U-test and Pearson's Chi-square test were used to compare the baseline characteristics between the two groups. Friedman test was conducted to see the within-group differences at different assessment occasions. If there was a significant difference within a group, post-hoc Wilcoxon signed-rank test would be carried out with Bonferroni correction. Mann–Whitney U-test was used to compare the between-group differences. For all participants in both groups, the mean age was 77.4 years and the percentage of males was 51.3%. As shown in Table 3, there were no significant differences in the baseline characteristics of participants between the experimental and control groups.

Primary outcomes

There was a significant increase in MFAC, EMS and LEFS in both groups from baseline to the second month, yet no significant difference between the groups was observed. The experimental group showed a significant increase in EMS and LEFS from baseline to the first month, whilst the control group showed a significant increase in LEFS only (Table 4). There were no significant differences in MFAC, EMS and LEFS among both groups at the first month and second month (Table 5).

Secondary outcomes

The experimental group showed better exercise adherence than the control group with significant difference in the first month. Although exercise adherence in the experimental group was still higher than the control group in the second month, there was no significant difference among the two groups (Table 5). Significant reduction in M-CSI was observed in both groups from baseline to the second month, yet no significant difference between groups was observed (Table 4).

Table 3. Baseline characteristics of the participants.

Participants' characteristics	Experimental group ($n = 19$)	Control group ($n = 20$)	p-Value (Mann–Whitney U)	p-Value (Chi-square)
Age (years), mean (SD)	75.8 (7.2)	79 (8.8)	0.204	
Gender, number of males (%)	12 (63.5)	8 (40)		0.148
Premorbid MFAC (SD)	6.8 (0.7)	6.7 (0.9)	0.813	
Types of fracture (%)				0.329
Femoral neck	10 (52.6)	12 (60)		
Inter-trochanteric	7 (36.8)	8 (40)		
Sub-trochanteric	2 (10.5)	0 (0)		
Types of operation (%)				0.276
Hemiarthroplasty	7 (36.8)	6 (30)		
Total hip replacement	0 (0)	3 (15)		
Dynamic hip screw	2 (10.5)	0 (0)		
Cannulated screw	2 (10.5)	2 (10)		
Intramedullary nail	8 (42.1)	9 (45)		
Hospitalization (days), mean (SD)	26.5 (3.9)	26.5 (4)	0.728	
Caregiver relationships (%)				0.418
Spouse	8 (42.1)	5 (25)		
Children	11 (57.9)	13 (65)		
Relative	0 (0)	1 (5)		
Domestic helper	0 (0)	1 (5)		
MFAC (SD)	5.5 (1)	5.3 (1.1)	0.550	
EMS (SD)	11.8 (5.2)	12 (3.5)	0.687	
LEFS (SD)	25.7 (15.9)	20.5 (7.9)	0.496	
M-CSI (SD)	6 (6.4)	4.8 (4.7)	0.879	

Note: Difference between groups by the Mann–Whitney U-test or Chi-square test.

Table 4. Comparison of outcome measures from baseline to one month and from baseline to two months.

	From baseline to one month (T_0–T_1)		Between-group	From baseline to two months (T_0–T_2)		Between-group
	Exp. ($n = 19$)	Con. ($n = 20$)	p-value	Exp. ($n = 19$)	Con. ($n = 20$)	p-value
MFAC			0.901			0.728
Median	1	0.5		1	1	
(IQR)	(0–1)	(0–1)		(0.5–2)	(0–1.5)	
p-Value	0.266	0.399		0.002*	0.008*	
EMS			0.351			0.647
Median	3	2		4	3	
(IQR)	(1–5.5)	(0.5–5)		(2–7)	(2.5–5)	
p-Value	0.004*	0.081		<0.001*	<0.001*	
LEFS			0.945			0.411
Median	16	13.5		18	22	
(IQR)	(10.5–22)	(8.5–21)		(13–34)	(19–33)	
p-Value	0.005*	0.01*		<0.001*	<0.001*	
M-CSI			0.184			0.531
Median	−1	−2		−2	−3	
(IQR)	(−3–0)	(−4.5–−1)		(−7–0)	(−6–−1)	
p-Value	0.370	0.002*		<0.001*	<0.001*	

Note: Exp.: Experimental group; Con.: control group; IQR: interquartile range; difference within groups by the post-hoc Wilcoxon signed-rank test; difference between groups by the Mann–Whitney U-test. *$p < 0.05$.

Table 5. Comparison of outcome measures at the baseline, first month and second month post-discharge.

	Baseline (T_0)		First month (T_1)		Between-group	Second month (T_2)		Friedman test: Between-group	Friedman test: *p*-value	
	Exp. ($n = 19$)	Con. ($n = 20$)	Exp. ($n = 19$)	Con. ($n = 20$)	*p*-value	Exp. ($n = 19$)	Con. ($n = 20$)	*p*-value	Exp. ($n = 19$)	Con. ($n = 20$)
MFAC (1–7)					0.22			0.11	<0.001*	0.001*
Median	6	5	6	6		6	6			
(IQR)	(5–6)	(4.5–6)	(6–6)	(5–6)		(6–7)	(6–6)			
EMS (0–20)					0.43			0.34	<0.001*	<0.001*
Median	14	13 15	15		17	17				
(IQR)	(8.5–15)	(10–14)	(14–17.5)	(14–16)		(15.5–19.5)	(16–18)			
LEFS (0–80)					0.34			0.43	<0.001*	<0.001*
Median	20	20	42	34.5		52	47.5			
(IQR)	(17.5–28)	(14–26)	(29.5–56)	(26–47.5)		(30–62.5)	(34.5–54)			
M-CSI (0–26)			0.19		0.51	<0.001*	<0.001*			
Median	4	3.5	1	0		0	0			
(IQR)	(0–11.5)	(1–7)	(0–5)	(0–3)		(0–1.5)	(0–1.5)			
Adherence (0–100%)					0.03*			0.09		
Median (%)			100	75		95	67.5			
(IQR) (%)			(78.5–100)	(54–100)		(76.5–100)	(50–100)			

Note: Exp.: Experimental group; Con.: control group; IQR: interquartile range; difference within groups by the Friedman test; difference between groups by the Mann–Whitney *U*-test. $^*p < 0.05$.

Discussion

The aim of this study was to investigate the efficacy of a home-based rehabilitation program using a mobile app on elderly patients with hip fracture and their caregivers. The results showed that these patients using the mobile app had better exercise adherence when compared to using pamphlets. There were two features in the mobile app, which pamphlets could not provide, that might lead to higher motivation of patients in performing exercises. First, a daily exercise reminder using text messages was delivered to patients' mobile devices. This feature was highly noticeable as the reminder activated with an audible alarm that drew the attention of the elderly patients. The patients could then see the motivational messages as soon as they picked up the mobile devices. Text messaging was already shown to be an effective intervention in improving exercise adherence among elderly population.[33] Second, video-guided exercise programs were embedded in the mobile app so that patients could perform exercise along with the videos. Compared to text and illustrations, video demonstrations were able to provide clear and timely instructions, such as close-up views and verbal cues for better exercise techniques. This mode of exercise instruction might improve patients' confidence in performing the home-based exercises without therapists' supervision. A study suggested that there was a positive correlation between self-efficacy and adherence rate to exercise interventions for hip fracture patients.[34] The effect of increased exercise adherence in patients using the mobile app was not existent in the second month. It might be due to a possibility that the patients began to disregard the daily exercise reminder, as the effect of motivational messages was known to diminish over time.[35] Future development of mobile app could investigate the possibility of sending feedback to the therapists when under-utilization of the mobile app was detected, so that they could arrange timely home visits for exercise re-enforcement.

Patients using either mobile app or exercise pamphlets had significant improvement in the primary outcomes (MFAC, EMS and LEFS) over two months, however, there were no additional benefits demonstrated in using the mobile app. It was thought that improvement in exercise adherence would lead to better physical performance. However, this relationship was not established in the study. A possible reason for this phenomenon might be that the exercises in the home-based

rehabilitation program focused on strengthening and functional movements only, but the primary outcomes also take into account the performance in other aspects such as walking speed, stairs walking and outdoor activities. The outbreak of COVID-19 during the study period might have reduced patients' motivation to participate in outdoor activities, thus limiting the potential in physical performance and self-efficacy. Furthermore, a follow-up period of two months might also not be sufficient to observe the effects of mobile apps on hip fracture patients. A study demonstrated that hip fracture patients could have physical improvements in a longer study period of six–nine months.[24] Further studies could be done to investigate the benefits of the mobile app on physical performance and self-efficacy over a longer period of time.

In this study, caregivers from both groups reported relatively low stress level. It was different from the observation in other studies that caregivers experienced a moderate-to-high level of stress after the hip fracture patients got discharged from the hospital.[11,36] It could be due to the fact that the participants in this study had a high baseline mobility status and did not require much assistance in transfer and ambulation at home. Hence, the provision of the mobile app for the experimental group or the educational pamphlet for the control group might not make a difference in reducing the stress of these caregivers. Further studies could be conducted to investigate the efficacy of the mobile app in reducing caregiver stress with dependent patients.

Caregiver support is key to ensure the successful implementation of home rehabilitation using the mobile app, especially for elderly patients who would have a higher chance to experience issues with modern technology. Several common technical issues were identified throughout the study. Many participants were unable to install the mobile app in their mobile devices alone, since they often forgot or did not have the record of username and password to authorize apps installation. Some participants thought that the mobile app could be installed in their mobile devices, however, they found that the operating system of the devices was neither iOS nor Android. The screen size of some participants' mobile devices was found to be too small that videos and texts within the app became unreadable, and their caregivers had to provide these participants with another mobile device. A recommendation to resolve the above issues is to establish a checklist of mobile device requirements and involve caregivers early before hospital discharge, so as to prevent any delays in home rehabilitation.

There were several limitations in this study. It was carried out during the outbreak of COVID-19 and the recruitment of participants was affected. The small sample size reduced the power to detect differences between the groups. A longer follow-up period would help understand the long-term effect of a home-based rehabilitation program on geriatric hip fracture patients. Ascertainment bias might occur since neither outcome assessors nor participants were blinded from knowing which treatment group they belonged to. Participants and their caregivers were involved in this study, therefore, the results could not be generalized to elderly patients who live alone.

Conclusions

Use of the mobile app improved exercise adherence, yet it did not improve physical performance, self-efficacy and reduce caregiver stress when compared to a standard home rehabilitation program for elderly patients with hip fracture. Further studies to investigate the benefits of mobile apps are required.

Conflict of Interest

The authors declare no conflicts of interest relevant to this work.

Funding/Support

The authors received no financial support for the research, authorship and/or publication of this paper.

Author Contributions

Conception and design of the study were carried out by Kin Ming Ken Lau, Fuk On Titanic Lau, Mun Cheung Herman Lau and Sheung Wai Law. Acquisition of data was carried out by Kui Ching Cheng and Tin Sing Keith Lau. Analysis and interpretation of data were carried out by Kui Ching Cheng and Andy S. K. Cheng. Drafting of the manuscript was carried out by Kui Ching Cheng, Kin Ming Ken Lau, Andy S. K. Cheng and Tin

Sing Keith Lau. Manuscript revision was carried out by Tin Sing Keith Lau, Fuk On Titanic Lau, Mun Cheung Herman Lau and Sheung Wai Law. All authors approved the final manuscript to be published.

References

1. Akesson K, Marsh D, Mitchell PJ, McLellan AR, Stenmark J, Pierroz DD, et al. Capture the fracture: a best practice framework and global campaign to break the fragility fracture cycle. Osteoporos Int 2013;24(8):2135–52.
2. Johnell O, Kanis J. Epidemiology of osteoporotic fractures. Osteoporos Int 2005;16(Suppl 2):S3–7.
3. Gullberg B, Johnell O, Kanis J. World-wide projections for hip fracture. Osteoporos Int 1997;7 (5):407–13.
4. Cooper C, Campion G, Melton3rd LJ. Hip fractures in the elderly: A world-wide projection. Osteoporos Int 1992;2(6):285–9.
5. Man LP, Ho WH, Wong SH. Excess mortality for operated geriatric hip fracture in Hong Kong. Hong Kong Med J 2016;22(1):6–10.
6. Shumway-Cook A, Ciol MA, Gruber W, Robinson C. Incidence of and risk factors for falls following hip fracture in community-dwelling older adults. Phys Ther 2005;85(7):648–55.
7. Portegijs E, Edgren J, Salpakoski A, Kallinen M, Rantanen T, Alen M, et al. Balance confidence was associated with mobility and balance performance in older people with fall-related hip fracture: a cross-sectional study. Arch Phys Med Rehabil 2012;93(12):2340–6.
8. Magaziner J, Simonsick EM, Kashner TM, Hebel JR, Kenzora JE. Predictors of functional recovery one year following hospital discharge for hip fracture: a prospective study. J Gerontol 1990;45(3): M101–7.
9. Magaziner J, Hawkes W, Hebel JR, Zimmerman SI, Fox KM, Dolan M, et al. Recovery from hip fracture in eight areas of function. J Gerontol A Biol Sci Med Sci 2000;55(9):M498–507.
10. Kristensen MT. Factors affecting functional prognosis of patients with hip fracture. Eur J Phys Rehabil Med 2011;47(2):257–64.
11. Ariza-Vega P, Ortiz-Piña M, Kristensen MT, Castellote-Caballero Y, Jiménez-Moleón JJ. High perceived caregiver burden for relatives of patients following hip fracture surgery. Disabil Rehabil 2017;41(3):311–8.
12. Pisters MF, Veenhof C, de Bakker DH, Schellevis FG, Dekker J. Behavioural graded activity results in better exercise adherence and more physical activity than usual care in people with osteoarthritis: a cluster-randomised trial. J Physiother 2010;56(1): 41–7.
13. Beinart NA, Goodchild CE, Weinman JA, Ayis S, Godfrey EL. Individual and intervention-related factors associated with adherence to home exercise in chronic low back pain: a systematic review. Spine J 2013;13(12):1940–50.
14. Friedrich M, Cermak T, Maderbacher P. The effect of brochure use versus therapist teaching on patients performing therapeutic exercise and on changes in impairment status. Phys Ther 1996;76 (10):1082–8.
15. Rhodes RE, Martin AD, Taunton JE, Rhodes EC, Donnelly M, Elliot J. Factors associated with exercise adherence among older adults: An individual perspective. Sports Med 1999;28(6):397–411.
16. Mohammed BMA, Abdelkhalek WSS, Hassanine ST, Amen AIF, Alghamdi W, Elmetwaly AAM. Effect of booklet-based education versus mobile-based education on women's knowledge and practices related lymphedema regarding post mastectomy exercise. Egypt J Health Care 2001;12(4): 1474–96.
17. Goto M, Takedani H, Haga N, Kubota M, Ishiyama M, Ito S, et al. Self-monitoring has potential for home exercise programmes in patients with haemophilia. Haemophilia 2014;20(2):e121–7.
18. Lambert TE, Harvey LA, Avdalis C, Chen LW, Jeyalingam S, Pratt CA, et al. An app with remote support achieves better adherence to home exercise programs than paper handouts in people with musculoskeletal conditions: a randomised trial. J Physiother 2017;63(3):161–7.
19. Cotter AP, Durant N, Agne AA, Cherrington AL. Internet interventions to support lifestyle modification for diabetes management: A systematic review of the evidence. J Diabetes Complications 2014;28(2):243–51.
20. Hutchesson MJ, Rollo ME, Krukowski R, Ells L, Harvey J, Morgan PJ, et al. eHealth interventions for the prevention and treatment of overweight and obesity in adults: a systematic review with meta-analysis. Obes Rev 2015;16(5):376–92.
21. Wildenbos GA, Peute LW, Jaspers MW. A framework for evaluating mHealth tools for Older Patients on Usability. Stud Health Technol Inform 2015;210:783–7.
22. Yeung PY, Wong LL, Chan CC, Leung JL, Yung CY. A validation study of the Hong Kong version of Montreal Cognitive Assessment (HK-MoCA) in Chinese older adults in Hong Kong. Hong Kong Med J 2014;20(6):504–10.
23. Orwig DL, Hochberg M, Yu-Yahiro J, Resnick B, Hawkes WG, Shardell M, et al. Delivery and outcomes of a yearlong home exercise program after hip fracture: a randomized controlled trial. Arch Intern Med 2011;171(4):323–31.
24. Latham NK, Harris BA, Bean JF, Heeren T, Goodyear C, Zawacki S, et al. Effect of a home-based

exercise program on functional recovery following rehabilitation after hip fracture: a randomized clinical trial. JAMA 2014;311(7):700–8.
25. American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription. 8th ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins, 2010.
26. Chau MWR, Chan SP, Wong YW, Lau MYP. Reliability and validity of the Modified Functional Ambulation Classification in patients with hip fracture. Hong Kong Physiother J 2013;31(1):41–4.
27. Smith R. Validation and reliability of the Elderly Mobility Scale. Physiotherapy 1994;80(11):744–7.
28. Prosser L, Canby A. Further validation of the Elderly Mobility Scale for measurement of mobility of hospitalized elderly people. Clin Rehabil 1997;11(4):338–43.
29. Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. Phys Ther 1999;79(4):371–83.
30. Hou WH, Yeh TS, Liang HW. Reliability and validity of the Taiwan Chinese version of the Lower Extremity Functional Scale. J Formos Med Assoc 2014;113(5):313–20.
31. Thornton M, Travis SS. Analysis of the reliability of the modified caregiver strain index. Gerontol B Psychol Sci Soc Sci 2003;58(2):S127–32.
32. Chan WCH, Chan CLF, Suen M. Validation of the Chinese version of the Modified Caregivers Strain Index among Hong Kong caregivers: An initiative of medical social workers. Health Soc Work 2013;38 (4):214–21.
33. Müller AM, Khoo S, Morris T. Text messaging for exercise promotion in older adults from an upper-middle-income country: Randomized controlled trial. J Med Internet Res 2016;18(1):e5.
34. Resnick B, D'Adamo C, Shardell M, Orwig D, Hawkes W, Hebel JR, et al. Adherence to an exercise intervention among older women post hip fracture. J Clin Sport Psychol 2008;2(1):41–56.
35. Figueroa CA, Deliu N, Chakraborty B, Modiri A, Xu J, Jay Williams J, et al.Daily motivational text messages to promote physical activity in university students: Results from a microrandomized trial. Ann Behav Med. 2022;56(2):212–8.
36. Lin PC, Lu CM. Hip fracture: family caregivers' burden and related factors for older people in Taiwan. J Clin Nurs 2005;14(6):719–26.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 42, No. 2 (2022) 111–124
DOI: 10.1142/S1013702522500111

Hong Kong Physiotherapy Journal
https://www.worldscientific.com/worldscinet/hkpj

Normative reference values and regression equations to predict the 6-minute walk distance in the Asian adult population aged 21–80 years

Meredith T. Yeung[1,*], Melissa Y. Chan[1], Katherin S. Huang[1,2], Tian Jie Chen[1], Cyprian P. Chia[1,3], Meihiko M. Fong[1], Cherilyn S. Ho[1], Derek T. Koh[1,4], Mitchell J. Neo[1] and Mark Tan[1]

[1]*Health and Social Sciences Cluster*
Singapore Institute of Technology, Singapore

[2]*Department of Physiotherapy*
Khoo Teck Puat Hospital, Singapore

[3]*Department of Physiotherapy*
Alexandra Hospital, Singapore

[4]*Department of Physiotherapy*
Sengkang Community Hospital, Singapore

[*]*Meredith.Yeung@SingaporeTech.edu.sg*

Received 4 February 2022; Accepted 3 May 2022; Published 24 June 2022

Summary at a glance: The 6-min walk test (6MWT) is a widely used field walking test. This study reports the normative reference values (NRV) of distance walked during 6MWT (6MWD) in healthy Singaporeans (aged 21–80) and updates the 6MWD reference equations. This information may facilitate the interpretation of the 6MWD in clinical populations.
Ethics approval: The Singapore Institute of Technology-Institutional Review Board (SIT-IRB Project Number: 2019099) approved this study to be carried out from June 2019 to January 2021. All participants gave written informed consent before data collection began.

[*]Corresponding author.

Copyright@2022, Hong Kong Physiotherapy Association. This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) License which permits use, distribution and reproduction, provided that the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Background: The six-minute walk test (6MWT) is a widely adopted submaximal field-walking test to evaluate functional exercise capacity. This validated test is a reliable, safe, inexpensive, and straightforward assessment tool commonly used as an outcome measure, using the distance walked (6MWD) as the primary outcome. An earlier study has established the normative reference values (NRV) and equation in healthy Singaporeans — however, the small sample size and narrow age range curb adequate representation of the adult population profile.
Objectives: This study aims to update the NRV and reference equations to predict the distance walked during 6MWT (6MWD) for healthy Singaporeans aged 21–80.
Methods: This cross-sectional study recruited community-dwelling healthy subjects aged 21–80 via convenience sampling. Each subject completed two trials of 6MWT according to the standard protocol. Primary outcome measures included 6MWD, pre-and post-test heart rate (HR), oxygen saturation, and blood pressure (BP).
Results: 172 healthy Singaporeans (females = 90, males = 82) participated. The overall mean 6MWD was 578.00 ± 75.38 metres. The age-stratified mean 6MWD ranged from 601.3 ± 71.79 metres (aged 21–39) to 519.02 ± 55.42 metres (aged 60–80). Age, gender, and percentage maximum HR predicted (%PredHRmax) were the most significant variables ($p < 0.001$). 6MWD reference equation $= 288.282$ (height, m) $+ 27.463 \times$ Gender (male $= 1$; female $= 0$) $+ 4.349$ (%predHRmax) $+ 1.191$ (HR reserve, bpm) $-185.431 - 1.343$ (age, years) $- 1.614$ (weight, kg), $R^2 = 58\%$. Applying equations from other studies to the Singaporean population resulted in an overestimation of the 6MWD.
Conclusion: This study updated the NRV and reference equations of 6MWD for healthy Singaporeans aged between 21–80 years. This update revises the local benchmarks of 6MWD in Singapore, a widely adopted outcome measure.

Keywords: 6-Minute walk test; exercise test; exercise capacity; outcome measures; reference values and equations.

Introduction

The six-minute walk test (6MWT) is a widely-used submaximal field walking test to evaluate functional exercise capacity.[1–5] The 6MWT is a safe, inexpensive, and straightforward test commonly used for assessment and as an outcome measure, using the distance walked (6MWD) as the primary outcome. It is valid and reliable in young and adult, healthy and diseased populations such as chronic obstructive pulmonary disease, pulmonary fibrosis, heart failure, and amputees.[6–11] As the test is self-paced and rests are allowed, it is considered an appropriate assessment of exercise capacity for individuals suffering from respiratory-related symptoms and the general geriatric population.[12–14] Thus, it has an advantage over other comprehensive cardiopulmonary exercise tests (CPET) in the clinical setting.

Judgement of the individual walking performance should be compared with that of a relevant population, and it requires the availability of normative references (norms) for the particular population.[1,15] Several studies have derived the normative reference values (NRV) and reference equations to estimate 6MWD specific to the respective local populations.[16–26] Factors, such as demographic and anthropometric profiles, clinical and physiological characteristics, have been reported as variables affecting the 6MWD.[3,4,8,16–19,21–24,27–31] Poh *et al.* further confirmed that the reference equations derived from the Caucasian population would overestimate the 6MWD in healthy Singaporeans.[22] Till date, the study by Poh *et al.* remains the only attempt to report the NRV and reference equation for healthy Singaporeans despite tremendous changes in the Singaporean population profile in the last 15 years.[32] Additionally, the small sample size ($n = 35$) and age range of 45–85 years in their study lacked representation of the healthy adult Singaporean population. Furthermore, the use of a 45 metres (m) corridor is uncommon in the land-scarce city-state. Thus, there is a need to update the NRV and derive reference equations that could be applicable to a broader age range.

Therefore, this study aims to do the following: (1) establish the NRV of 6MWD in the healthy

Singaporean population aged 21–80 years; (2) determine the correlations of variables that could influence the 6MWD; (3) establish the 6MWD reference equation applicable to healthy adult Singaporeans; (4) evaluate the age-matched comparisons of the Singapore data with published studies.

Methods

Study design

This cross-sectional study was conducted via convenience sampling at various community centres in Singapore from June 2019 to January 2021. Approval of the study was obtained from the University Institutional Review Board (Project number: 2019099). Personal particulars and written informed consent were obtained from every subject before the commencement of the study.

Subjects

We recruited community-dwelling healthy Singaporean subjects aged 21–80 via convenience sampling from various districts in Singapore. A single mean and standard deviation (SD) method was used to determine the sample size. With reference to the previous study[21,26] that had reported the NRV of 6MWD, we predicted a small-medium effect size [Cohen's $d = 0.312$ [95% confidence interval (95%CI) 0.283–0.893]] with a minimum sample size of 142 subjects based on 95% confidence; expected population SD of 60 metres (m) and a precision of 10.[33] We allowed for a possible attrition rate of 20%, and thus a minimum of 170 subjects were needed to distribute across the age range. All subjects completed the Physical Activity Readiness Questionnaire for Everyone (PAR-Q$^+$).[34] before data collection to determine suitability. Inclusion criteria were: individuals between the ages of 21–80 years as of testing day; able to understand simple English; able to walk a minimum of six minutes independently; and clearance with the PAR-Q$^+$. Subjects were excluded if they had a history of chronic disease(s), such as cardiovascular, metabolic, respiratory, neuromuscular or musculoskeletal conditions that could affect the ability to perform physical exercises or cause changes to functional capacity; gait abnormality, such as the use of walking aid or leg length discrepancy ≥ 2 cm, measured from the anterior superior iliac spine to lateral malleolus, as such discrepancy is associated with prevalent, incident symptomatic, and progressive knee osteoarthritis;[35] current smoking history or any smoking history within the last 12 months; resting heart rate (HR) > 100 beats per minute (bpm) or < 50 bpm; resting systolic BP > 150 or < 90 mmHg, resting diastolic BP > 100 or < 50 mmHg; resting oxygen saturation (SpO_2) $< 95\%$, abnormal lung function, which is defined as forced expiratory volume in 1 second (FEV_1) $\leq 80\%$ or forced vital capacity (FVC) $\leq 80\%$ or $FEV_1/FVC \leq 70\%$.[36] These values were referenced from and kept consistent with other NRV reports on 6MWD for subsequent comparisons between the derived HR variables.[21,22,26]

Data collection

Researchers who were proficient in conducting the 6MWT performed the data collection. All subjects were instructed to wear comfortable clothes and walking shoes. The data obtained from subjects before undergoing 6MWT included PAR-Q$^+$, age, gender, height, weight, smoking history, medical history and medication use, HR, BP, SpO_2, dyspnoea score with the modified Borg's dyspnoea scale,[37] and Borg's rating of perceived exertion (RPE).[38] Body-mass index (BMI) calculation followed the standard formula.[39] The maximum predicted HR (PredHRmax) was calculated using $[208 \times (0.7 \times \text{Age})]$,[40] while heart rate reserve (HRR) was calculated with (PredHRmax – Resting HR). The percentage of maximum predicted HR (%predictHRmax) was derived using [(measured highest HR during 6MWT $\div$ PredHRmax) $\times 100\%$]. A calibrated standard portable spirometer (Spirolab, MIR) was used to measure lung functions. Each subject had to complete at least three measures according to the standard guidelines,[41] with the highest values of the FEV_1, FVC, and FEV_1/FVC considered for inclusion and exclusion criteria.

Six-minute walk test

The 6MWT was conducted along a 30-metre indoor walkway according to the guidelines of the American Thoracic Society (ATS).[1] Subjects walked back and forth and turned around two cones that marked the ends of the walkway. Subjects were instructed to "walk as far as possible in 6 minutes" and that they could slow down or rest if

necessary. At the end of every minute, the researchers gave standardised encouragement according to the ATS guidelines[1] and recorded heart rate and oxygen saturation with a portable pulse oximeter (Nellcor™ PM10N), dyspnoea and RPE scores.[37,38] At the end of 6 min, the 6 MWD was recorded. Two trials of 6 MWT were conducted on the same day to observe for learning effect,[11] with a minimum 30-minute rest interval between the two trials to ensure sufficient recovery and that HR, BP, and SpO_2 returned to baseline before the second trial. The longer distance walked between the two trials was used for all subsequent data analyses.

Statistical analysis

GraphPad Prism Version 8.4.3 (686) (GraphPad Software, San Diego, California, USA) was used to perform the statistical analysis. The level of significance was set at $p < 0.05$. Demographic and anthropometric data of subjects were examined for normal distribution using the Shapiro–Wilk test. Descriptive statistics were used to analyse central tendency, data spread, and the dataset's position via means, SD, 95% confidence interval (95%CI), percentage and percentiles accordingly. Chi-square Goodness of Fit test was used to compare the ethnicity ratio between the subject profiles and the Singaporean population.[42] Test–retest reliability was examined using the interclass correlation coefficient (ICC). Kruskal–Wallis test was used to analyse non-normal continuous variables by age groups, and Mann–Whitney U test to compare the variables by gender. Pearson's correlation coefficients assessed the correlation between variables. Linear regression was applied to establish the reference equation for 6MWD. To compare the age-matched Singapore data with the published reference equations, the measured 6MWD from this study was age-matched to each study and compared against the same age range of the calculated distance derived from the predictive regression formulae of the eleven identified reports.[16–26] The comparisons between measured and predicted 6MWD were analysed using paired *t*-tests.

RESULTS

Subject characteristics and 6MWD

One hundred and seventy-two subjects (82 males; 90 females) were included for data analysis from 175 subjects recruited, with three excluded based on exclusion criteria. Tables 1 and 2 present the demographics and characteristics of the subjects. The Chi-Square Goodness of Fit test showed a statistically significant difference ($p = 0.01$) in the subject ethnicity profile compared to the population profile of Singapore 2021,[42] with a higher percentage of Chinese subjects. The overall mean 6 MWD was 578.00 ± 75.38 m, where males walked an average of 600.32 ± 77.29 m while females averaged 557.62 ± 67.84 m ($p < 0.001$) (Table 1). The 6 MWD decreased progressively with age and ranged from 601.30 ± 71.79 m (age 21–39) to 519.02 ± 55.42 m (age 60–80) (Table 2). Test–retest reliability was excellent (ICC = 0.90). The highest HR achieved was 133.4 ± 21.48 bpm from the 21–39 year group; while the %PredHRmax ranged from $70.07 \pm 11.16\%$ (age 21–39) to 76.40 $\pm 11.76\%$ ($p = 0.008$) (Table 2).

Variable associations with 6MWD and regression equations

The relationships between 6MWD, demographics, anthropometric and physiological variables are presented in Table 3 and Fig. 1. These parameters are commonly collected before and during the 6MWT assessment as suggested by guidelines,[1] illustrating their readiness and potential usefulness in estimating and benchmarking the outcome measurement. 6MWD was significantly associated with age, gender, height, weight, average leg length, HRR and %PredHRmax (all $p < 0.05$). Notably, the correlation of average leg length ($r = 0.36$, $p < 0.001$) and height ($r = 0.38$, $p < 0.001$) to 6MWD was similar; also, a high correlation ($r = 0.83$, $p < 0.05$) was established between average leg length and height, thus only height was included in the linear regression analysis.

Stepwise linear regression analysis revealed that age, height, weight, gender, %PredHRmax, HRR were independent contributors to 6MWD (Table 4). Pre-6MWT available variables such as age and height, or age and gender alone explained 23–24% of the variance, respectively. The better reference equation was as follows: 6MWD (m) = 622.64 + 35.03 × Gender (male = 1; female = 0) −1.65 (age, years). However, with the addition of the post-test available variables, the percentage of the variance increased to 58%; and the reference equation was as follows: 6MWD (m) = 288.28 (height, m) + 27.46 × Gender (male = 1; female = 0) + 4.35 (%predHRmax) + 1.19

Table 1. Subjects' characteristics and measured variables during the 6MWT (by gender).

Characteristics	Total	Male	Female	*p*-value
Subjects, n	172	82	90	—
Age (years)	37.1 ± 18.5	34.7 ± 17.7	39.3 ± 19.0	0.10
Ethnicity, *n* (%)				0.01*
Chinese	158 (91.9)	77 (44.8)	81 (47.0)	—
Malay	5 (2.9)	3 (1.7)	2 (1.2)	—
Indian	9 (5.2)	2 (1.2)	7 (4.1)	—
Height (m)	1.66 ± 0.09	1.72 ± 0.07	1.60 ± 0.06	< 0.001
Weight (kg)	64.69 ± 13.90	72.44 ± 14.35	57.62 ± 8.82	< 0.001
BMI (kg/m^2)	23.46 ± 3.80	24.47 ± 3.85	22.54 ± 3.53	< 0.001
Average leg length (m)	0.87 ± 0.05	0.89 ± 0.05	0.84 ± 0.04	< 0.001
HR measurements				
Resting HR (bpm)	85.75 ± 14.49	82.87 ± 13.13	88.38 ± 15.22	< 0.001
HRR (bpm)	96.29 ± 18.29	100.87 ± 18.47	92.12 ± 17.19	< 0.001
Highest HR (bpm)	130.54 ± 20.58	129.35 ± 21.16	131.62 ± 20.09	0.472
HRchange (bpm)	44.79 ± 19.19	46.49 ± 19.87	43.24 ± 18.53	0.270
%PredHRmax	71.91 ± 11.32	70.62 ± 11.73	73.08 ± 10.85	0.154
6MWD (m)				
6MWD1	562.13 ± 74.74	586.21 ± 76.03	540.20 ± 66.72	< 0.001
95% CI	550.90 to 573.40	569.51 to 602.92	526.2 to 554.20	
6MWD2	570.33 ± 74.91	590.45 ± 77.22	552.00 ± 68.13	< 0.001
95% CI	559.1 to 581.60	573.48 to 607.41	537.70 to 566.30	
6MWD2 – 6MWD1	8.20 ± 32.90	4.24 ± 34.46	11.80 ± 31.17	0.13
95% CI	3.24 to 13.15	−3.34 to 11.08	5.28 to 18.33	
Best of 2 trials	578.00 ± 75.38	600.32 ± 77.29	557.62 ± 67.84	< 0.001
95% CI	566.60 to 589.30	583.3 to 617.3	543.4 to 571.8	

Notes: Values are expressed as *mean* ± *Standard Deviation* (*SD*),
$p < 0.05$ represents a significant value.
m: metres; **kg**: kilogram; **bpm**: beats per minute; **BMI**: Body mass index; **6MWT**: 6-minute walk test; **6MWD**: 6-minute walk distance; **HR**: Heart Rate; **Highest HR**: Highest heart rate achieved during 6MWT; **HRR**: heart rate reserve; **HRchange**: Difference between HighestHR heart rate and resting heart rate; **%PredHRmax**: peak HR achieved during 6MWD expressed as %predicted maximum HR with predicted HRmax as [208−(0.7×Age)]; **95% CI**: 95% Confidence Interval; **6MWD1**: 1st trial of 6MWD; **6MWD2**: 2nd trial of 6MWD; *subject ethnicity profile compared with Singapore population profile.

Table 2. Subjects' characteristics and measured variables during the 6MWT (by age).

	Age			
Characteristics	21–39 years	40–59 years	60–80 years	*p*-value
Subjects, *n*	115	20	37	
Ethnicity, *n* (%)				
Chinese	106 (61.6)	15 (8.7)	37 (21.5)	
Malay	5 (2.9)	0 (0)	0 (0)	
Indian	4 (2.4)	5 (2.9)	0 (0)	
Height (m)	1.68 ± 0.08	1.60 ± 0.08	1.60 ± 0.08	< 0.001
Weight (kg)	67.23 ± 14.80	58.33 ± 8.59	59.52 ± 10.27	0.003
BMI (kg/m^2)	23.64 ± 3.94	22.84 ± 2.65	23.20 ± 3.86	0.812
Average leg length (m)	0.88 ± 0.05	0.84 ± 0.03	0.83 ± 0.05	< 0.001
HR measurements				
Resting HR (bpm)	86.58 ± 14.51	85.00 ± 14.89	83.42 ± 14.35	0.740
HRR (bpm)	103.76 ± 14.46	85.98 ± 14.50	76.98 ± 14.42	< 0.001
Highest HR (bpm)	133.4 ± 21.48	130.3 ± 13.37	121.78 ± 18.73	< 0.001

Table 2. (*Continued*)

Characteristics	Age			*p*-value
	21–39 years	40–59 years	60–80 years	
HRchange (bpm)	46.88 ± 19.70	44.35 ± 13.03	38.53 ± 19.46	< 0.001
%PredHRmax	70.07 ± 11.16	74.97 ± 8.33	76.40 ± 11.76	0.008
6MWD (m)				
6MWD1	582.9 ± 73.91	538.60 ± 52.97	510.30 ± 57.57	< 0.001
95% CI	569.2 to 596.5	513.80 to 563.4	491.1 to 529.50	
6MWD2	593.1 ± 71.40	547.1 ± 62.03	512.00 ± 54.96	< 0.001
95% CI	579.90 to 606.30	518.00 to 576.10	493.70 to 530.30	
6MWD2 – 6MWD1	10.25 ± 36.06	8.45 ± 31.27	1.68 ± 21.23	0.116
95% CI	3.59 to 16.91	−6.185 to 23.08	−5.40 to 8.76	
Best of 2 trials	601.30 ± 71.79	553.00 ± 60.64	519.00 ± 55.42	< 0.001
95% CI	588.00 to 614.60	524.60 to 581.30	500.50 to 537.50	

Notes: Values are expressed as *mean* ± *Standard Deviation* (*SD*),
$p < 0.05$ represents a significant value.
m: metres; **kg**: kilogram; **bpm**: beats per minute; **BMI**: Body mass index; **6MWT**: 6-minute walk test; **6MWD**: 6-minute walk distance; **HR**: Heart Rate; **Highest HR**: Highest heart rate achieved during 6MWT; **HRR**: heart rate reserve; **HRchange**: Difference between HighestHR heart rate and resting heart rate; **%PredHRmax**: peak HR achieved during 6MWD expressed as %predicted maximum HR with predicted HRmax as [208−(0.7×Age)]; **95% CI**: 95% Confidence Interval; **6MWD1**: 1st trial of 6MWD;**6MWD2**: 2nd trial of 6MWD; *subject ethnicity profile compared with Singapore population profile.

Table 3. Univariate correlation coefficients (r) for 6MWD and subject variables ($n = 172$).

Variable	r	95% CI	*p*-value
Age (years)	−0.44	−0.5497 to −0.3063	< 0.001*
Gender	0.28	0.1400 to 0.4157	< 0.001*
Height (m)	0.38	0.2389 to 0.4969	< 0.001*
Weight (kg)	0.17	0.02043 to 0.3113	0.02*
BMI (kg/m^2)	−0.44	−0.1925 to 0.1062	0.57
Average leg length (m)	0.36	0.2187 to 0.4806	< 0.001
Resting HR (bpm)	0.06	−0.09541 to 0.2030	0.47
HRR (bpm)	0.27	0.1207 to 0.3993	< 0.001*
%PredHRmax	0.36	0.2204 to 0.4821	< 0.001*

Notes: *$p < 0.05$ represents a significant value; CI 95%: 95% confidence intervnal; **m**: metres; **kg**: kilogram; Body mass index; **bpm**: beats per minute; **6MWD**: 6-minute walk distance; **HR**: Heart Rate; **HRR**: heart rate reserve; **%PredHRmax**: the percentage that highest HR achieved out of HRmax.

(HRR, bpm) −185.43−1.34 (age, years) −1.61 (weight, kg).

Comparisons with the 6MWD estimated using the previously published equations

Table 5 presents the age-matched measured data from this study compared to the age-matched 6MWD calculated using the predictive formulae from 11 previous similar studies.[16–26] The distance walked by the participants in this study was shorter than the calculated distances in 9 out of the 11 studies compared ($p < 0.001$), demonstrating the over-estimations with their predictive formulae,[16–18,20,22–26] with the exception of the study by Fernandes *et al.*[19] where our participants walked a significantly greater distance (91.06±

65.95) m, $p < 0.001$. The comparison made with Ngai *et al.* was statistically insignificant ($p = 0.22$).[21]

DISCUSSION

This study established an updated NRV and formulated the reference equations for 6MWD of healthy Singaporeans adults aged 21–80 years. This study reported that the overall mean 6MWD was 578.00 ± 75.38 m. Male subjects were found to walk a significantly longer distance than female subjects (600.32 ± 77.29 m versus 557.62 ± 67.84 m; $p < 0.001$), while the mean 6MWD decreased progressively with the advancement of age, from 601.30 ± 71.79 m (age 21–39 years) to 519.02 ± 55.42 m (age 60–80 years). The high ICC for repeated tests demonstrated good test–retest reliability and is consistent with previous studies.[21,24] The 6MWD is influenced by the learning effect[1,11,17,43] as the second test distance was consistently higher than the first test, but no more than 2%, which is well within the reported ranges for the healthy and diseased populations.[1,17,20,21,24,44] Two standardised tests were performed on the same day to avoid biases caused by the learning effect in this study, with the best of the two used in the analysis.

Age, gender, height, weight, leg length, HRR and %predHRmax influenced the 6MWD significantly in this study (Table 3), while BMI and resting HR were found to be statistically insignificant. Our subjects achieved 71.91 ± 11.32 %PredHRmax during the 6MWT (Table 1), indicating moderate-intensity effort, consistent with previous studies.[17,22]

Age and gender accounted for 24% of the 6MWD variance, while the influence of %predictHRmax explained 25.4% of the variance. This suggests that

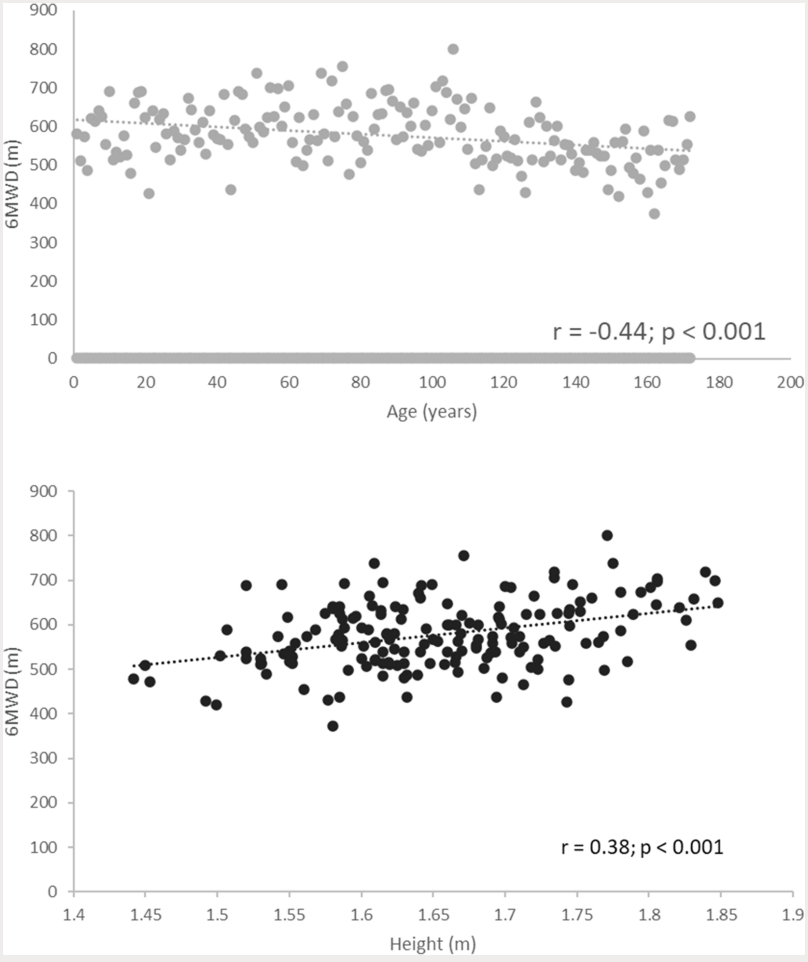
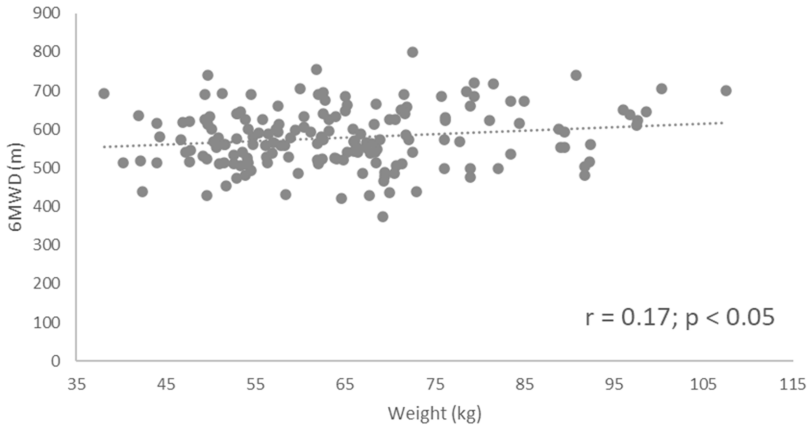
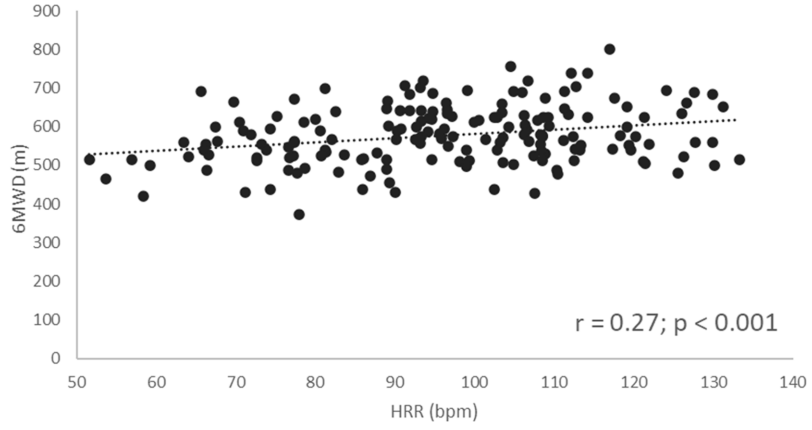


Fig. 1. The relationships between the 6MWD and Age, Height, Weight, HRR, and %PredHRmax; $p < 0.05$ represents the significance of an independent variable in predicting 6MWD.





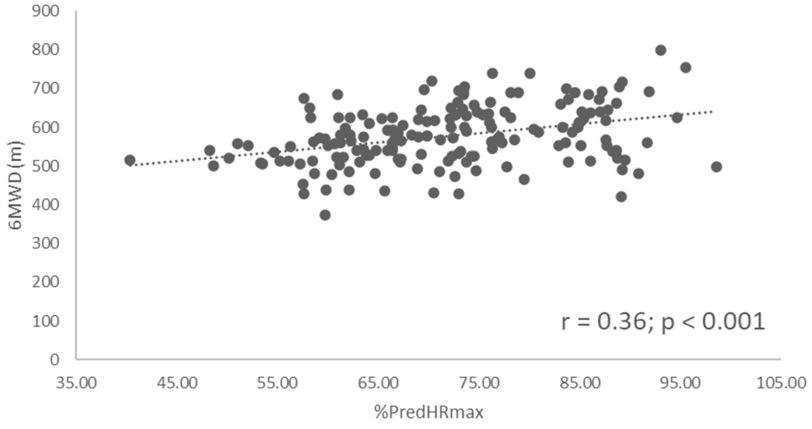


Fig. 1. (*Continued*)

age, gender, and %predictHRmax are the three most significant factors influencing the 6MWD. This report established two reference equations; one comprising simple pre-test variables with age and gender ($R^2 = 24\%$), indicating its predictive yet pragmatic application with obvious variables for any straightforward interpretation of the 6MWD. This simplified reference equation can also be particularly relevant to individuals on chronotropic agents, as such chronotropic effects may affect the regularity of the heart rate and rhythm, thus affecting the %predictHRmax. In contrast, the second reference equation comprised more variables and established a higher percentage of variance ($R^2 = 58\%$), which would be helpful for clinicians to benchmark the 6MWT performance of patients

Table 4. Linear regression models for predicting 6MWD.

Model	Independent variables	R^2	R^2 change	P (model)	Unstandardised coefficients		p (independent variables)
					B (95% CI)	SE	
	Pre-6MWT variables						
1	(Constant)	0.19	—	< 0.001	643.74 (620.77 to 666.70)	11.63	< 0.001
	Age				−1.77 (−2.33 to −1.22)	0.28	< 0.001
2	(Constant)	0.23	0.04	< 0.001	279.93 (255.79 to 304.07)	11.02	< 0.001
	Age				−1.39 (−1.94 to −0.87)	0.3	< 0.001
	Height				211.21 (62.55 to 359.88)	63.66	< 0.001
3	(Constant)	0.24	0.05	< 0.001	622.64 (597.31 to 647.96)	12.83	< 0.001
	Age				−1.65 (−2.2 to −1.11)	0.27	< 0.001
	Gender				35.03 (15.00 to 55.07)	10.15	< 0.001
	All 6MWT variables						
4	(Constant)	0.42	0.23	< 0.001	425.03 (368.83 to 481.24)	28.47	< 0.001
	Age				−2.27 (−2.75 to −1.78)	0.25	< 0.001
	%predHRmax				3.30 (2.50 to 4.09)	0.4	< 0.001
5	(Constant)	0.494	0.073	< 0.001	390.10 (335.59 to 444.60)	27.61	< 0.001
	Age				−2.15 (−2.61 to −1.69)	0.23	< 0.001
	%predHRmax				3.45 (2.70 to 4.20)	0.38	< 0.001
	Gender				41.24 (24.74 to 57.73)	8.36	< 0.001
6	(Constant)	0.524	0.03	< 0.001	233.08 (124.50 to 341.67)	55	< 0.001
	Age				−1.62 (−2.17 to −1.08)	0.28	< 0.001
	%predHRmax				4.01 (3.21 to 4.81)	0.41	< 0.001
	Gender				36.03 (19.69 to 52.38)	8.28	< 0.001
	HRR				1.03 (0.41 to 1.66)	0.32	< 0.001
7	(Constant)	0.542	0.018	< 0.001	274.38 (162.96 to 385.80)	56.43	< 0.001
	Age				−1.73 (−2.27 to −1.19)	0.28	< 0.001
	%predHRmax				4.17 (3.37 to 5.97)	0.4	< 0.001
	Gender				48.88 (30.01 to 67.75)	9.56	< 0.001
	HRR				1.06 (0.45 to 1.68)	0.31	0.001
	Weight				−0.89 (−1.58 to −0.21)	0.35	0.011
8	(Constant)	0.58	0.037	< 0.001	−185.43 (−445.62 to 74.76)	131.78	0.161
	Age				−1.34 (−1.90 to −0.78)	0.28	< 0.001
	%predHRmax				4.35 (3.58 to 5.12)	0.39	< 0.001
	Gender				27.46 (6.23 to 48.70)	10.75	0.012
	HRR				1.19 (0.60 to 1.78)	0.3	< 0.001
	Weight				−1.61 (−2.37 to −0.86)	0.38	< 0.001
	Height				288.28 (139.62 to 436.95)	75.3	< 0.001

Notes: $*p < 0.05$ represents a significant value; **6MWD**: 6-minute walk distance; **B**: unstandardised regression coefficient; **95% CI**: 95% Confidence Interval; **SE**: standard error; **HR**: Heart Rate; **%PredHRmax**: peak HR achieved during 6MWD expressed as %predicted maximum HR with predicted HRmax as [208 − (0.7 × Age)]; Gender: (1 = male, 0 = female); **HRR**: heart rate reserve (beats per minute).

or clients. The influence of age and gender on 6MWD has been well reported in earlier studies about 6MWD reference equations.[15,17–19,23,24,28] The negative association of age with 6MWD could be due to a decline in physical fitness, as it has been found that ageing results in muscle loss, reduction in $VO_{2\,max}$,[45–47] decreased stride length, and altered gait. Similarly, it is well established that females generally have lower cardiovascular and muscular fitness than males.[48–53] The influence of %predictHRmax, as a measure of cardiac response during a submaximal self-paced field walking test accounts for a significant proportion of the variability in test performance, suggests that cardiac response can be used as a surrogate measure of the exercise effort during the test.[20] Although including %predHRmax

Table 5. Age-matched comparison between measured 6MWD from this study and predicted 6MWD from reference equations of published studies.

Study (Predictive Equation)	Country	Measured 6MWD (m) mean ± SD	Predicted 6MWD (m) mean ± SD	Difference (m) (Measured− Predicted) mean ± SD	*p*-value
Poh *et al.* (2006) 45–85 years 6MWD (m) = 5.50 (%predHRmax) +6.94 (height, cm)− 4.49 (age, year) − 3.51 (weight, kg) − 473.27	Singapore	527.73 ± 59.15	560.33 ± 74.17	−32.60 ± 69.42	< 0.001
Camarri *et al.* (2006) 55–75 years 6MWD (m) = 216.90 + 4.12 (height, cm) − 1.75 (age, years) −1.15 (weight, kg) − 34.04 (gender : males = 0; females = 1)	Australia	530.71 ± 59.66	675.17 ± 44.11	−144.46 ± 56.44	< 0.001
Chetta *et al.* (2006) 20–50 years 6MWD (m) = 518.85 + 1.25 (height, cm) − 2.82 (age, year) −39.07 (gender : males = 0; females = 1)	Italy	597.98 ± 18.63	636.52 ± 32.37	−38.54 ± 18.04	< 0.001
Jenkin *et al.* (2009) 45–85 years Males, 6MWD (m) = 748 − 6.32 (age, years) + 0.64 (height, cm) + 2.69 (%pred HRmax); Females, 6MWD (m) = 541 − 3.81 (age, years) + 1.80 (height, cm) − 6.92 (BMI) + 2.41 (%pred HRmax)	Western Australia	529.08 ± 19.99	618.24 ± 51.66	−87.90 ± 30.10	< 0.001
Ben Saad *et al.* (2009) 40–85 years 6MWD (m) = 720.50 − 160.27 × gender (gender: males = 0; females = 1) −5.14 × (age, years) − 2.23 × (weight, kg) +271.98 × (height, m)	North Africa	533.21 ± 22.98	601.36 ± 89.74	−68.26 ± 74.02	< 0.001
Casanova *et al.* (2011) 40–80 years 6MWD = 361 − 4 (age, years) + 2 (height, cm) + 3 (HRmax/HRmax % *pred*) −1.5 (weight, kg) − 30 (gender : males = 0; females = 1)	North & South America	535.95 ± 48.87	552.49 ± 46.65	−16.28 ± 21.12	< 0.001
Kim *et al.* (2014) 22–59 years 6MWD (m) = 105.7 + 2.99 × (height, cm)	Korea	592.04 ± 26.34	605.70 ± 25.31	−13.08 ± 46.80	< 0.001
Ngai *et al.* (2014) 55–85 years 6MWD (m) = 722.35 − 5.11 × (age, years) + 2.19 ×(%predHRmax) − 41.31 × gender (gender: males = 0; females = 1)	Hong Kong	527.75 ± 19.61	535.95 ± 40.87	−8.88 ± 44.49	= 0.22
Fernandes *et al.* (2016) 25–75 years 6MWD (m) = 553.289 − 2.11 × (age, years) + 45.323 ×gender (gender: male = 1, female = 0)	West India	566.53 ± 78.59	475.47 ± 48.91	91.06 ± 63.95	< 0.001
Zou *et al.* (2017) 18–30 years Female: 6MWD (m) = −0.458 + (difference in heart rate ×1.113) + 3.494 × (height, cm); Male: 6MWD (m) = −11.394 + (difference in heart rate ×0.692) + 3.659 × (height, cm)	China	594.13 ± 72.14	618.65 ± 29.82	−24.51 ± 68.54	< 0.001
Oliveira *et al.* (2019) 18–70 years 6MWD = 721.7 − 1.6 × (Age, years) − 4.0 × BMI +0.9 × ΔHR × 58.4 × gender (gender: male = 1; female = 0)	Portugal	581.01 ± 35.31	639.17 ± 48.14	−58.16 ± 28.74	< 0.001

Notes: Values expressed as *mean* ± *Standard Deviation* (*SD*).
*$p < 0.05$ represents a significant difference between age-matched subjects measured vs predicted 6MWD. **6MWD**: 6-minute walk distance.

increases the variance of the reference equation, there are several limitations in the clinical setting as the magnitude of HR change could be influenced by external factors such as the use of chronotropic agents. As such, we consider it necessary to report both equations so that clinicians can choose to apply the appropriate equation to those who may experience chronotropic impairments during the test, e.g., patients who have heart failure or beta-blocker use.

The reported 6MWD reference equations explained up to 58% of the variance, similar to other studies' findings,[15,17–19,23,24,28] despite the different age groups of the subjects. Eleven published studies (Table 4) were chosen to perform the age-matched comparison between measured 6MWD from this study and predicted 6MWD using their reference equations. Of the 11 studies being compared, nine overestimated the 6MWD[16–18,20,22–24] and one study underestimated the distance.[19] Only the 6MWD regression formula reported by Ngai *et al.* is not statistically significant compared to the Singaporean data.[21] The previously published reference equations did not reliably predict the 6MWD in our population. We found significant differences between measured 6MWD and 6MWD predicted from the published equations and could postulate two reasons. First, the published predictive formulae were derived from the NRV of the respective populations. Notably, Camarri *et al.*[6] reported that Australians covered a total of 655 ± 51 m (male: 685 ± 49 m; female: 628 ± 59 m) during the 6MWT; Ben Saad *et al.*[16] reported the average 6MWD for the West Africans was 624 ± 111 m (male: 711 ± 81 m; female: 511 ± 75 m), while the reference values for Western Indians were 483 ± 67.91 m (male: 512.38 ± 67.84 m; female 457.29 ± 56.75 m).[19] These reported values vary considerably from the NRV of our study. Naturally, the eventual calculated predictive 6MWD matched for the Singaporean anthropometric variables would differ significantly from the current findings, despite the consideration of age-matching. Second, the effect of %PredHRmax on 6MWD is well-established from previous studies.[19,21–26] and in this current report (25.4% of the variance). The range of %PredHRmax in the other studies differed substantially from $51.9 \pm 8.8\%$,[19] $63.0 \pm 10.5\%$,[26] $69.0 \pm 8.0\%$,[23] $80.0 \pm 10.0\%$,[21] and $87.0 \pm 13.0\%$[24]; however, this is congruent with the nature of a self-paced submaximal field walking test. Submaximal and self-paced testing relies heavily on self-perception, concepts which are reflected well in the variance in %PredHRmax among the different populations. Hence, we suggest this is a possible explanation as to why only the Ngai *et al.* predicted distance (total 6MWD: 563.0 ± 62.0 m; %PredHRmax: $80.0 \pm 10\%$) was statistically insignificant to our current results, while similar reports from China (total 6MWD: 502.0 ± 73.0 m; %PredHRmax: $69.0 \pm 8.0\%$)[23] and South Korean (total 6MWD: 598.5 ± 57.92 m; %PredHRmax: $63.0 \pm 10.48\%$)[26] were not. The implications of this, particularly in adults with chronic diseases, may include considerable errors regarding the level of disability and unrealistic expectations of the outcome measure. This justifies the use of our local specific reference equations and confirms the ATS recommendation to continue establishing updated regional reference equations.[1]

Some limitations in this study should be considered. First, the profile of existing subjects is statistically different from the overall Singaporean population profile despite presenting the three major ethnicities living in Singapore, with an over-representation of the Chinese ethnicity. Second, 67% of the subjects from this study were in the 21–39 age group, while the 40–59 and 60–80 age groups combined contributed to the remaining 33% of the overall sample size. This possibly skewed the results to the 20–39 age group with an over-representation of younger adults. Third, this study did not explicitly collect additional psychological data, such as depression, balance confidence, or fear of falling that might potentially influence the 6MWD, despite such information being screened via the PAR-Q$^+$. For example, the General Health Questions section of PAR-Q$^+$ screened for the loss of balance due to dizziness in the last 12 months in Question 3, while Question 6 from the Follow-up Questions section checked for mental health problems such as depression or anxiety.[34] Future studies should consider the inclusion of such information. Finally, there was also a 42% variance that could not be explained with the existing data. Future studies should establish missing variables that could account for the remaining variance.

Conclusions

This study updated the NRV and reference equations of 6MWD for healthy Singaporean adults aged 21–80 years. Age, gender, height, weight, HRR and %predHRmax were significantly correlated to 6MWD. Applying equations from other studies to the Singaporean population resulted in an overestimation of the 6MWD. The reference equation and NRV should be beneficial in establishing performance benchmarks to guide intervention and rehabilitation. Future follow-up studies should consider exploring reasons for the unexplained variance in our equation.

Acknowledgments

The authors would like to thank the Health and Social Sciences Cluster of the Singapore Institute of Technology for the support and all the subjects who participated in the study. Ms Jade Soh for statistical support throughout the study.

Conflict of Interest

The authors declare no potential conflicts of interest for the research, authorship, and/or publication of this paper.

Funding/Support

This study received funding from the Health and Social Sciences Cluster, Singapore Institute of Technology, Singapore.

Author Contributions

Meredith T. Yeung conceptualised and designed the study, analysed and interpreted the data, prepared and revised the manuscript critically for important intellectual content, and read and approved the final manuscript. Melissa Y. Chan and Katherin S. Huang conceptualised and designed the study and drafted and revised the manuscript critically for important intellectual content. Tian Jie Chen, Cyprian P. Chia, Meihiko M. Fong, Cherilyn S. Ho, Derek T. Koh, Mitchell J. Neo, and Mark Tan collected, analysed and interpreted the data. All authors approved the final manuscript.

References

1. ATS. ATS statement: Guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166:111–7.
2. Cooper KH. A means of assessing maximal oxygen intake. Correlation between field and treadmill testing. JAMA 1968;203:201–4.
3. Enright PL. The six-minute walk test. Respir Care 2003;48:783–5.
4. Troosters T, Gosselink R, Decramer M. Six minute walking distance in healthy elderly subjects. Eur Respir J 1999;14:270–4.
5. Camarri B, Eastwood PR, Cecins NM, Thompson PJ, Jenkins S. Six minute walk distance in healthy subjects aged 55–75 years. Respir Med 2006;100:658–65.
6. Cox PD, Frengopoulos CA, Hunter SW, Sealy CM, Deathe AB, Payne MWC. Impact of course configuration on 6-minute walk test performance of people with lower extremity amputations. Physiother Can 2017;69:197–203.
7. Demers C, McKelvie RS, Negassa A, Yusuf S. Reliability, validity, and responsiveness of the six-minute walk test in patients with heart failure. Am Heart J 2001;142:698–703.
8. du Bois RM, Weycker D, Albera C et al. Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference. Am J Respir Crit Care Med 2011;183:1231–7.
9. Li AM, Yin J, Yu CC et al. The six-minute walk test in healthy children: Reliability and validity. Eur Respir J 2005;25:1057–60.
10. Mänttäri, Suni, J., Sievänen, H., Husu, P., Vähä-Ypyä, H., Valkeinen, H., Tokola, K. and Vasankari, T. Six-minute walk test: A tool for predicting maximal aerobic power (VO2 max) in healthy adults. Clinical Physiology and Functional Imaging 2018;38(6):1038–1045. https://doi.org/10.1111/cpf.12525.
11. Singh SJ, Puhan MA, Andrianopoulos V et al. An official systematic review of the European Respiratory Society/American Thoracic Society: Measurement properties of field walking tests in chronic respiratory disease. Eur Respir J 2014;44:1447–78.
12. Pulmonary rehabilitation: Joint ACCP/AACVPR evidence-based guidelines. ACCP/AACVPR pulmonary rehabilitation guidelines panel. American College of Chest Physicians. American Association of Cardiovascular and Pulmonary Rehabilitation. Chest 1997;112:1363–96.
13. *American Association of Cardiovascular and Pulmonary Rehabilitation. Guidelines for Pulmonary Rehabilitation Programs.* 4th ed.: Human Kinetics, 2011, Chapter 4: pp. 31–48.
14. Yohannes AM, Connolly MJ. Pulmonary rehabilitation programmes in the UK: A national representative survey. Clin Rehabil 2004;18:444–9.
15. Rothstein JM, Echternach JL. Primer on measurement: An introductory guide to measurement issues, featuring the American Physical Therapy Association's standards for tests and measurements in physical therapy practice. Alexandria, Va: American Physical Therapy Association, 1993.
16. Ben Saad H, Prefaut C, Tabka Z et al. 6-Minute walk distance in healthy North Africans older than 40 years: Influence of parity. Respir Med 2009;103:74–84.
17. Casanova C, Celli BR, Barria P et al. The 6-min walk distance in healthy subjects: Reference standards from seven countries. Eur Respir J 2011;37:150–6.
18. Chetta A, Zanini A, Pisi G, et al. Reference values for the 6-min walk test in healthy subjects 20-50 years old. Respir Med 2006;100:1573–8.
19. Fernandes L, Mesquita AM, Vadala R, Dias A. Reference equation for six minute walk test in

healthy Western India Population. J Clin Diagn Res 2016;10:CC01–4.
20. Jenkins S, Cecins N, Camarri B, Williams C, Thompson P, Eastwood P. Regression equations to predict 6-minute walk distance in middle-aged and elderly adults. Physiother Theory Pract 2009;25:516–22.
21. Ngai SPC, Jones AYM, Jenkins SC. Regression equations to predict 6-minute walk distance in Chinese adults aged 55–85 years. Hong Kong Physiother J 2014;32(2):58–64.
22. Poh H, Eastwood PR, Cecins NM, Ho KT, Jenkins SC. Six-minute walk distance in healthy Singaporean adults cannot be predicted using reference equations derived from Caucasian populations. Respirology 2006;11:211–6.
23. Zou H, Zhang J, Chen X et al. Reference Equations for the six-minute walk distance in the healthy Chinese Han population, aged 18-30 years. BMC Pulm Med 2017;17:119.
24. Camarri B, Eastwood PR, Cecins NM, Thompson PJ, Jenkins S. Six minute walk distance in healthy subjects aged 55-75 years. Respir Med 2006;100:658–65.
25. Oliveira MJ, Marçôa R, Moutinho J et al. Reference equations for the 6-minute walk distance in healthy Portuguese subjects 18–70 years old. Pulmonology 2019;25:83–9.
26. Kim AL, Kwon JC, Park I et al. Reference equations for the six-minute walk distance in healthy Korean Adults, aged 22–59 years. Tuberc Respir Dis 2014;76:269.
27. Burr JF, Bredin SS, Faktor MD, Warburton DE. The 6-min walk test as a predictor of objectively measured aerobic fitness in healthy working-aged adults. Phys Sportsmed 2011;39:133–9.
28. Cote CG, Pinto- Plata V, Kasprzyk K, Dordelly LJ, Celli BR. The 6-min walk distance, peak oxygen uptake, and mortality in COPD. Chest 2007;132:1778–85.
29. Giannitsi S, Bougiakli M, Bechlioulis A, Kotsia A, Michalis LK, Naka KK. 6-min walking test: A useful tool in the management of heart failure patients. Ther Adv Cardiovasc Dis 2019;131753944719870084.
30. Lee R, Chan YH, Wong J, Lau D, Ng K. The 6-min walk test predicts clinical outcome in Asian patients with chronic congestive heart failure on contemporary medical therapy: A study of the multiracial population in Singapore. Int J Cardiol 2007;119:168–75.
31. Puhan MA, Mador MJ, Held U, Goldstein R, Guyatt GH, Schunemann HJ. Interpretation of treatment changes in 6-min walk distance in patients with COPD. Eur Respir J 2008;32:637–43.
32. Singstat. Population and Population Structure In: Singapore DoS, ed. *Population*. https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/visualising-data/population-dashboard: Department of Statistics Singapore, 2020.
33. Ramsey F, Schafer D. The Statistical Sleuth: A Course in Methods of Data Analysis. 2nd ed. Duxbury Thomson Learning, 2002.
34. Warburton DE, Jamnik VK, Bredin SS, Gledhill N. The physical activity readiness questionnaire for everyone (PAR-Q+) and electronic physical activity readiness medical examination (ePARmed-X+). Health Fitness J Canada 2011;4:3–17.
35. Harvey WF, Yang M, Cooke TD et al. Association of leg-length inequality with knee osteoarthritis: A cohort study. Ann Internal Med 2010;152:287–95.
36. Pellegrino R, Viegi G, Brusasco V et al. Interpretative strategies for lung function tests. Eur Resp J 2005;26:948–68.
37. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14:377–81.
38. Borg GA. Perceived exertion. Exerc Sport Sci Rev 1974;2:131–53.
39. Deurenberg P, Weststrate JA and Seidell JC. Body mass index as a measure of body fatness: Age- and sex-specific prediction formulas. Br J Nutr 1991;65:105–14.
40. Tanaka H, Monahan KD and Seals DR. Age-predicted maximal heart rate revisited. J Am Coll Cardiol 2001;37:153–6.
41. Gardner RM and Crapo RO. New recommendations for standardising spirometry. Chest 1988;93:226–7.
42. Population and Population Structure. In: Environment PA, ed. https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/latest-data: Singapore Department of Statistics, 2021.
43. Holland AE, Spruit MA, Troosters T et al. An official European respiratory society/American thoracic society technical standard: Field walking tests in chronic respiratory disease. Eur Respir J 2014;44:1428–46.
44. Jenkins S, Cecins NM. Six-minute walk test in pulmonary rehabilitation: Do all patients need a practice test?Respirology 2010;15:1192–6.
45. Fleg JL, Lakatta EG. Role of muscle loss in the age-associated reduction in VO2 max. J Appl Physiol (1985) 1988;65:1147–51.
46. Dao T, Green AE, Kim YA et al. Sarcopenia and muscle aging: A brief overview. Endocrinol Metab (Seoul) 2020;35:716–32.
47. Ligibel JA, Schmitz KH, Berger NA. Sarcopenia in aging, obesity, and cancer. Transl Cancer Res 2020;9:5760–71.

48. Agarwal B, Shah M, Andhare N, Mullerpatan R. Incremental shuttle walk test: Reference values and predictive equation for healthy Indian adults. Lung India 2016;33:36–41.
49. Itaki M, Kozu R, Tanaka K et al. Reference equation for the incremental shuttle walk test in Japanese adults. Respir Investig 2018;56:497–502.
50. Jurgensen SP, Antunes LC, Tanni SE et al. The incremental shuttle walk test in older Brazilian adults. Respiration 2011;81:223–8.
51. Kenney WL, Wilmore J, Costill D. Physiology of Sport and Exercise. 6th ed. Human Kinetics, 2015.
52. Probst VS, Hernandes NA, Teixeira DC et al. Reference values for the incremental shuttle walking test. Respir Med 2012;106:243–8.
53. Harrison SL, Greening NJ, Houchen-Wolloff L et al. Age-specific normal values for the incremental shuttle walk test in a healthy British population. J Cardiopulm Rehabil Prev 2013;33:309–13.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 42, No. 2 (2022) 125–136
DOI: 10.1142/S1013702522500123

Hong Kong Physiotherapy Journal
https://www.worldscientific.com/worldscinet/hkpj

Validity and feasibility of using a seated push-up test among community-dwelling older adults

Puttipong Poncumhak[1,2,3,4], Supaporn Phadungkit[1,2,†], Pakwipa Chokphukiao[1,2], Roongnapa Intaruk[1,2], Pipatana Amatachaya[3,5] and Sugalya Amatachaya[1,2,*]

[1]*School of Physical Therapy, Faculty of Associated Medical Sciences Khon Kaen University, Khon Kaen, Thailand*

[2]*Improvement of Physical Performance and Quality of Life (IPQ) Research Group Khon Kaen University, Khon Kaen, Thailand*

[3]*Department of Physical Therapy, School of Allied Health Sciences University of Phayao, Phayao, Thailand*

[4]*Unit of Excellent of Physical Fitness and Exercise University of Phayao, Phayao, Thailand*

[5]*Department of Mechanical Engineering, Faculty of Engineering and Architecture Rajamangala University of Technology Isan, Nakhon Ratchasima, Thailand*

samata@kku.ac.th

Received 12 February 2022; Accepted 23 May 2022; Published 22 July 2022

Background: Older individuals face a high risk of mobility and body composition decline, which can affect their independence. In light of a current uncertain healthcare situation created by the coronavirus (COVID-19) pandemic, healthcare paradigm has been shifted with increased demand for a practical measure to promote standard home healthcare services for all individuals, including older adults.
Objective: This study explored the feasibility and validity of seated push-up tests (SPUTs) as clinical measures to reflect the body composition, muscle strength, and mobility among community-dwelling older individuals, aged $\geq$ 65 years ($n = 82$).

*Corresponding author.
†This author contributed equally with the first author, thus she is a co-first author.

Copyright@2022, Hong Kong Physiotherapy Association. This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) License which permits use, distribution and reproduction, provided that the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Methods: Participants were cross-sectionally assessed using SPUTs with various demanding forms, including the 1-time SPUT (1SPUT) along with its upper limb loading SPUT (ULL-SPUT), 5-time SPUT (5SPUT), 10-time SPUT (10SPUT), and 1-min SPUT (1minSPUT) and standard measures.
Results: Participants who passed and failed a 1SPUT showed significant differences in the outcomes of all standard measures ($p < 0.05$). The ULL-SPUT significantly correlated to all body composition, muscle strength, and mobility ($r = 0.247$–0.785; $p < 0.05$). Outcomes of 1minSPUT significantly correlated with muscle strength and mobility outcomes ($r = 0.306$–0.526; $p < 0.05$). Participants reported no adverse effects following the SPUTs.
Conclusion: The findings suggest the use of the 1SPUT, ULL-SPUT, and 1minSPUT as practical measures to reflect the body composition, muscle strength, and mobility of older individuals, according to their functional levels. The tests may especially clinically benefit those with lower limb limitations and those in settings with limited space and equipment.

Keywords: Body composition; clinical measure; endurance; mobility; muscle strength.

Introduction

The physiological decline accompanying aging occurs throughout the body's systems and results in a deterioration of physical function, particularly in the lower extremities. Consequently, approximately one-quarter of older people experience basic mobility limitations such as difficulties with walking, stair climbing, and rising from a chair, affecting their independence.[1–3] The lower limb impairments also distort the use of existing mobility measures that commonly involve lower limb functions such as the 10-m walk test, 6-min walk test, and sit-to-stand test. As a result, considerable effort has been directed towards understanding and attenuating age-related functional decline of the lower extremities. However, clinical measures involving upper limb functions, which might be applied instead among these individuals, have received relatively less attention.[4,5]

Existing evidence reports the use of an upper limb measure—namely, the handgrip test (HG)—to reflect total body strength, total skeletal muscle mass (SMM; $r = 0.49$; $p < 0.01$), and many other crucial aspects for older adults, such as health status and functional decline (odds ratio [OR] = 0.88).[6–8] However, the HG requires a specialised machine to assess the distal muscles of the tested arm in an open-kinetic chain manner. Such characteristics may affect the sensitivity of the outcomes in detecting problems among older individuals, as well as the clinical applications of the measurement in various community- and home-based settings. The present researchers hypothesised that the application of a practical measure involving several upper limb muscles working in a closed-kinetic chain manner may offer another clinical measure to detect common problems among older individuals. Such a test may be particularly beneficial for those with lower limb impairments and in settings with limited space or in which a specialised HG dynamometer is not available, especially in a current healthcare-paradigm shift with the need for standard home healthcare services due to coronavirus (COVID-19) pandemic.

A seated push-up test (SPUT) is a practical measure that can be executed on a chair or bed using push-up boards or wooden boxes. However, the task is very demanding and challenging for the upper limb and upper trunk muscles, as they must be able to exert enough muscle force and joint torque to lift the body upward by both arms and maintain body balance at the shoulder joints.[9,10] Wiyanad *et al.*[11] recently reported the ability of the upper limb loading during a seated push-up test (ULL-SPUT) to reflect body composition among individuals with a spinal cord injury. Our preliminary study also found an association between ULL-SPUT and the SMM of older individuals, particularly when the test is performed in a ring sitting position.[12] However, this preliminary study[12] investigated only the simplest form of SPUT—namely, the 1-time SPUT (1SPUT) along with its ULL-SPUT—in 40 well-functioning older adults. In addition, body composition was assessed only in terms of the SMM, using bioelectrical impedance analysis; the outcomes obtained might contain errors due to many factors, including electrode placement, environmental factors, and participant preparation.[13]

Based on the concept of global physiological change throughout the body systems and the closed-association of the musculoskeletal system, the present researchers hypothesised that various forms of SPUTs could be applied in older individuals with different functional levels and that their outcomes would reflect many aspects necessary for the independence of these individuals, depending upon the particular characteristics of the tests. Therefore, this study assessed the discriminative and concurrent validity, as well as the feasibility, of various types of SPUTs—including the 1SPUT along with its ULL-SPUT, 5-time SPUT (5SPUT), 10-time SPUT (10SPUT), and 1-min SPUT (1minSPUT)—as compared to the results of standard measures for body composition, muscle strength, and mobility among community-dwelling older adults.

Methods

Participants

This observational study was conducted among community-dwelling individuals aged 65 years and older with a body mass index (BMI) of between 18.5 kg/m^2 and 29.9 kg/m^2. The eligible participants needed to have the ability to stand up independently, walk with or without a walking device, and understand the instructions for the tests in this study. Individuals were excluded from the study if they had any signs or symptoms that might affect their participation in the study, such as uncontrolled medical conditions (e.g., hypertension or heart disease); pain in the musculoskeletal system that might affect outcomes of the study, such as a rotator cuff injury; and a history of shoulder or upper limb problems that limits their ability to perform SPUTs (i.e., a pain score of more than 5 out of 10 on a visual analog scale). All participants signed written informed consent forms that were approved by the Institutional Ethics Committee for Human Research (HE 611600). The estimated minimum sample size for this study was 82 participants, when $R_0 = 0.0$ and the lowest R_1 from a pilot study of 0.31 ($n = 40$), with 90% power and an alpha value of 0.05.[14]

Research protocols

The eligible participants were interviewed and assessed for their demographics, including age, gender, height, bodyweight, vital signs, underlying diseases, and walking device used (if any). Then, the participants were assessed for their ability to perform SPUTs, and standard measures for the body composition, muscle strength, and mobility of the older individuals. Details of the tests are explained below.

Seated push-up tests: Many forms of SPUTs are clinically available, including untimed and timed SPUTs.[9,10] With the aim to report the feasibility of using SPUTs among older individuals, this study applied simple SPUTs in increasingly demanding forms, including the 1SPUT along with its ULL-SPUT, 5SPUT, 10SPUT, and 1minSPUT, according to the participant's ability to complete the tests (with no pressure to complete all SPUTs if they were unable). Before and after the tests, the participants engaged in a warm-up and stretching session to reduce the risk of musculoskeletal injury, which can occur after completing such demanding measures. The various SPUTs were completed as follows.

The 1SPUT was executed using push-up loading devices at the size of standard clinical push-up boards (18 cm in height) to quantify the ULL-SPUT. The devices were developed from digital load cells (Model L6E3-C, 50 kg-3G, with the standard calibration method based on UKAS LAB 14: 2006; mini-patent application number 2103001612).[11,12] After calibration, the tools were accurate up to < 0.1 kg, with a measurement uncertainty of $\pm$ 0.1 kg. Participants were in a ring sitting position and placed their hands on the push-up loading devices slightly anterior to their hips (Fig. 1). Then, they pushed both hands against the devices, lifted the body from the floor while slightly bending the trunk forward and depressing both scapulars, and gradually bent the elbows to sit down on the floor.[9,11,12] The test was repeated over three trials, with a sufficient rest period between the trials. Outcomes of the test were recorded in terms as either *pass* or *fail*; a *pass* was defined as the ability to lift the body from the floor successfully in at least two of the three trials; if not, the outcome was regarded as a *fail*.[12] In addition, the average data of the maximum ULL-SPUT over the three trials, which was automatically generated by the push-up loading devices, was recorded.[11,12]

Participants who failed a 1SPUT terminated the SPUTs, and were assessed using standard measures. Participants who passed continued the timed-based SPUTs, including the 5SPUT,

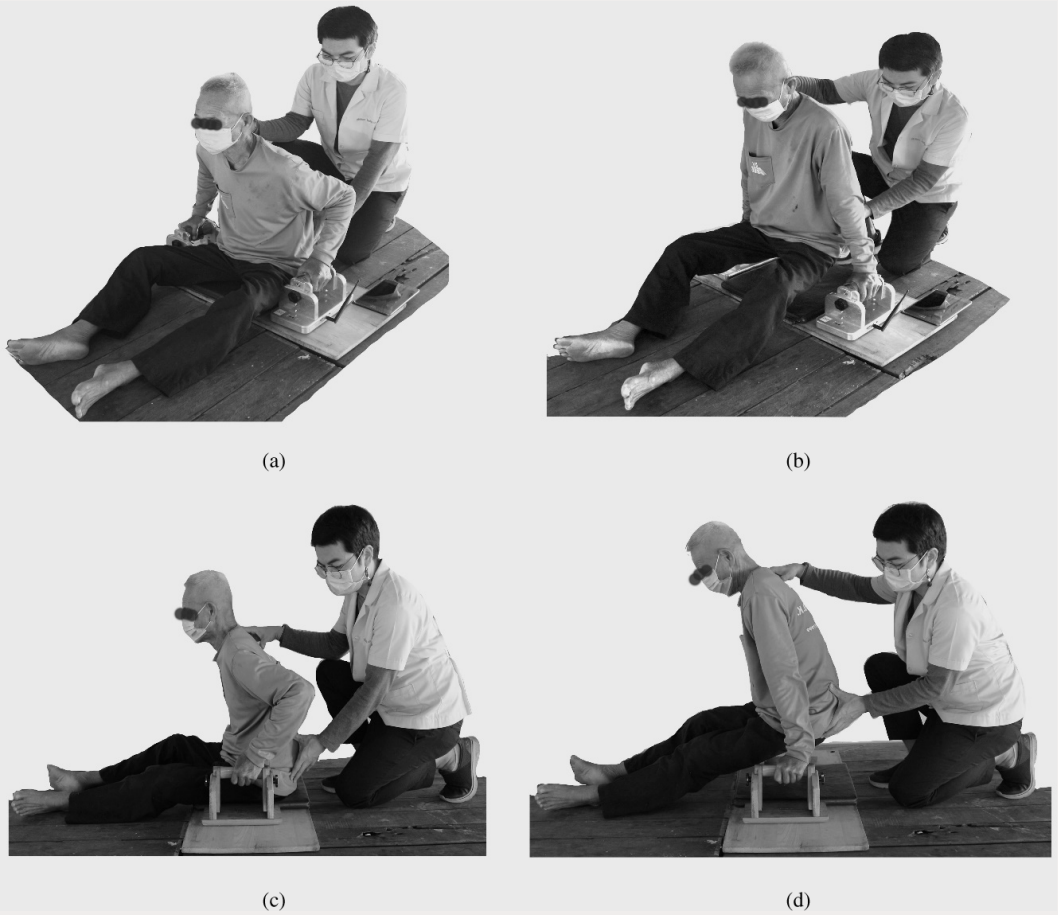
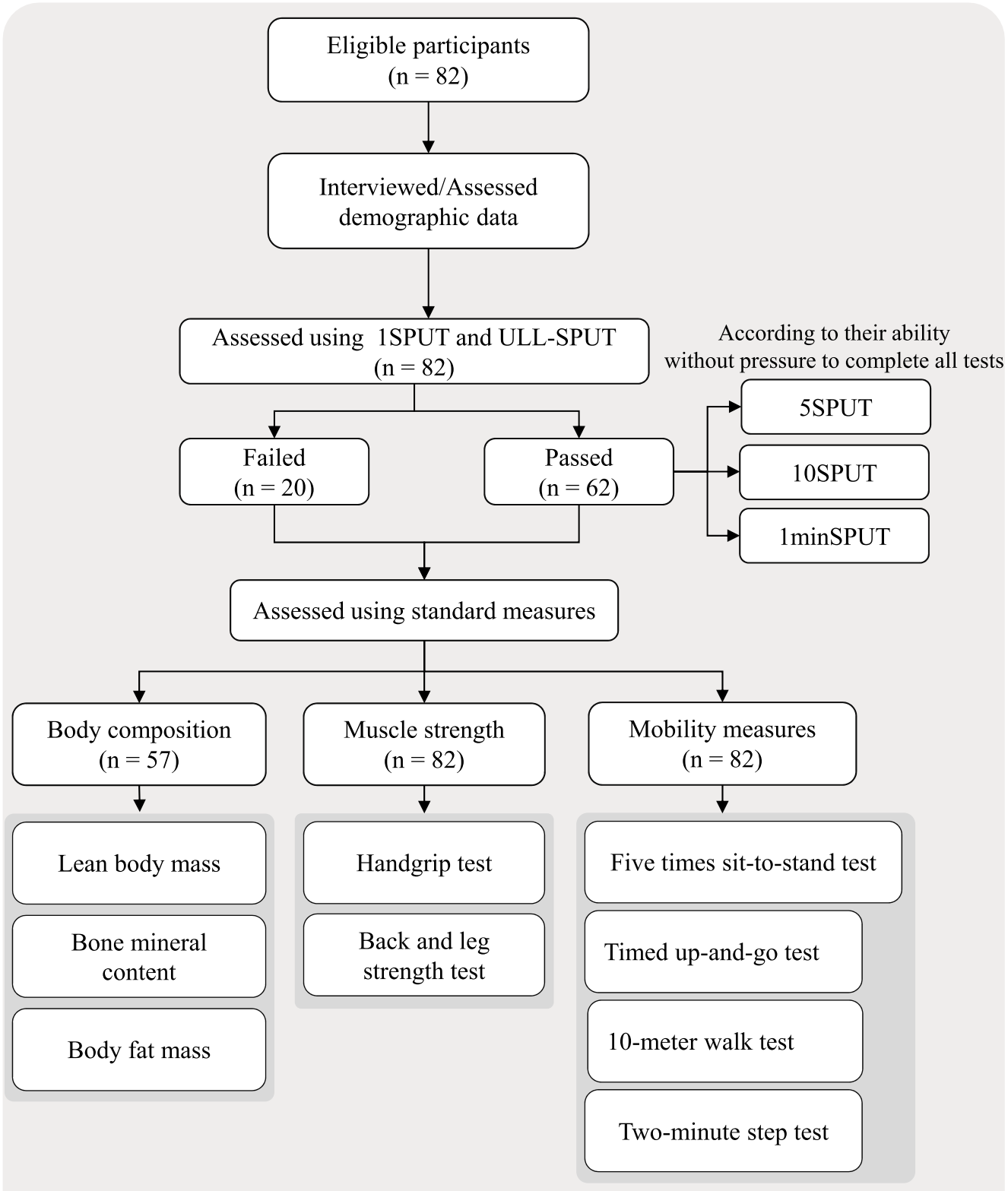


Fig. 1. Testing protocols of seated push-up tests. (A) Starting position with push-up loading devices. (B) Position while lifting the body from the surface with push-up loading devices. (C) Starting position with push-up boards. (D) Position while lifting the body from the surface with push-up boards.

10SPUT, and 1minSPUT—using standard clinical push-up boards and a starting position similar to that used for the 1SPUT. For the 5SPUT and 10SPUT, the participants were timed on their ability to complete 5 and 10 SPUT repetitions in the fastest and safe manner from the instruction "start" until the participant's buttock touched the floor on the last repetition. The average time over the three trials was recorded. For the 1minSPUT, the participants were assessed for the maximum number of SPUT repetitions they could do in one minute over one trial. During the test, they could take a period of rest as required and continue the test as soon as they could; otherwise, they terminated the test if they were unable to continue.

The SPUTs were assessed on a hard and level surface by an experienced rater (intraclass correlation coefficients [ICCs] = 0.932–1; $p < 0.001$). The participants could take a period of sufficient rest between the trials and the tests as required (at least a minute). The number of participants who could complete each method of SPUT was recorded, along with the adverse events (if any), such as musculoskeletal pain, chest pain, or accidental events, for the consideration on the feasibility of the SPUTs.

Standard measures: The participants were assessed using standard measures to indicate their body composition (i.e., lean body mass [LBM], bone mineral content [BMC], and body fat mass), muscle strength, and mobility (Table 1) by an experienced assessor in a random order. The participants could take a period of rest between the tests and the trials as required (or at least 30 s), in order to minimise the learning effects and fatigue that might occur due to the sequence of the tests. They were fastened with a lightweight safety belt around their waist so that the assessor could provide efficient assistance if needed. Details of the standard measures are described in Table 1.[15–34]



Notes: SPUT = seated push up test, ULL-SPUT = upper limb loading during a seated push up test

Fig. 2. Participation flowchart.

Table 1. Details of the standard measures for body compositions, muscle strength, and mobility.

Outcome measure		Aims of the assessments/ psychometric properties	Details of the measurement
Body compositions	Dual-energy X-ray absorptiometry (DXA)	Body compositions (lean body mass [LBM], bone mass content [BMC], and body fat mass [FM]).[16] Inter- and intra-tester reliability (ICC = 0.97–0.98).[17]	Participants were in a supine position on a DXA table with the arms placed on their sides, the legs extended, and the toes facing upward according to the standard protocols recommended by GE-Healthcare.[18] The data of body compositions, including LBM, BMC, and body fat mass, were automatically generated by the machine in kilograms.
Muscle strength tests	Handgrip test	Upper limb muscle strength, physical frailty, and disability in older individuals. Excellent test-retest reliability (ICC = 0.912–0.954).[19]	Participants were in a sitting position and squeezed the handle as much as they could over three trials per hand with a sufficient rest interval (at least 30 s) between the trials, and the maximum force was recorded.[20]
	Back and leg strength test	The back and leg extensor muscle strength. Excellent intra-tester reliability (ICC = 0.97).[21]	Participants stood on the base of a back-leg-chest dynamometer with the chain adjusted according to their height and the protocols for the trunk or leg extensor muscles explained previously. Then, they pulled the chain with their maximum force, holding it for 3 s. A rest period between the trials was allowed as needed (at least 30 s). The average force in the three trials for each method was recorded in kilograms.[22]
Mobility assessments	Five times sit-to-stand test	Functional lower extremity muscle strength and dynamic balance control while changing postures.[15,23] Test-retest reliability (ICC = 0.81; SEM = 0.9 s).[23,24] MDC_{95} = 2.5 s.[24] Concurrent validity with the TUG test; $r = 0.64$, $p < 0.001$.[24]	Participants were timed their ability to complete five chair-rise cycles at the fastest possible safe speed without using their arms. The average time over the three trials was reported.[25]
	Timed up-and-go test	Mobility, dynamic balance control, and risk of falling in older individuals.[26,27] Test-retest reliability (ICC = 0.97).[28] Inter-tester reliability (ICC = 0.99).[26]	Participants were timed their ability of standing up from a standard armrest chair, walking at the fastest and safe speed around a traffic cone at 3 m from the chair, returning, walking back to sit back down on the chair. The average time over the three trials was reported.[25]
	10-m walk test	Overall quality of gait, community participation, health condition, morbidity, and mortality rates.[29] Test-retest reliability (ICC = 0.92) and MDC = 0.22 m/s.[30]	Participants were timed their ability of walking at a comfortable pace along the middle 4 m of the total 10-m walkway. The average time over the three trials was then converted to a walking speed.[31,32]

Table 1. (*Continued*)

Outcome measure	Aims of the assessments/ psychometric properties	Details of the measurement
Two-minute step test	Functional capacity and physical endurance.[33] Test-retest reliability (ICC = 0.90) and concurrent validity with 1-mile walking time ($r = 0.73$).[34]	Participants raised their knee to a mid-thigh level, that is, the mid distance between the iliac crest and the patella, marking the point on the wall. The total number of steps in place, that is, the number of times the right knee reached the target level in 2 min over one trial was recorded.[33]

Notes: ICC: intraclass correlation coefficient; SEM: standard error of measurement; MDC: minimal detectable change.

Statistical analysis

Descriptive statistics were used to describe the participants' characteristics and the findings of the study. The independent samples *t*-test and Mann–Whitney *U*-test were used to compare the data between the participants who passed and failed the 1SPUT for the data with a normal and non-normal distribution, respectively (i.e., the discriminative validity). The Pearson correlation coefficient was used to analyse the correlations of continuous data between the SPUTs and standard measures (i.e., the concurrent validity). The correlation level was interpreted as being very low or negligible ($r = 0$–0.30), low ($r = 0.30$–0.50), moderate ($r = 0.50$–0.70), high or strong ($r = 0.70$–0.90), or excellent ($r = 0.90$–1.00).[14] Therefore, the closer the correlation coefficient was to 1, regardless of the direction, the stronger was the existing association, indicating a linear relationship between the SPUT data and the standard measures. The level of statistical significance was set at $p < 0.05$.

Results

In total, 82 individuals, with an average age of 74 years and a normal BMI, completed the study. Most participants were well-functioning females, physically active, and able to perform daily activities independently, without mobility devices ($n = 72$; 88%; Table 2). Of all, 57 participants were assessed for their body composition because this variable was additionally included after the initiation of the study.

All 82 participants could complete a 1SPUT; 62 participants (75.6%) passed the test and proceeded to be assessed with other forms of the SPUTs, while the rest of them ($n = 20$) failed in the test and continued with standard measures. Most participants who failed were female ($n = 18$; 90%), and approximately one-third of them ($n = 7$; 35%) used a single cane for daily movement. The average ULL-SPUT of the participants who passed was 85% of their bodyweight and for those who failed was 71% of their bodyweight ($p < 0.001$; Table 3).

Table 2. Personal data of all participants and of those who passed and failed a 1-seated push up test (1SPUT).

Variable	Total (n=82)	Fail ($n = 20$)	Pass ($n = 62$)	*P*-value
Age[a] (years)	74.6 ± 6.5 (73.1–76)	76.5 ± 7.4 (73–79.9)	74 ± 6.1 (72.4–75.5)	0.120
Gender[b] (female)	49 (61)	18 (90)	29 (46.7)	0.002*
Bodyweight[a] (kg)	55.6 ± 9.9 (53.5–57.8)	55.9 ± 11.9 (50.3–61.5)	55.6 ± 9.3 (53.2–57.9)	0.893
Body height[a] (m)	1.5 ± 0.1 (1.5–1.6)	1.5 ± 0.1 (1.5–1.6)	1.6 ± 0.1 (1.5–1.6)	0.906
Body mass index[a] (kg/m^2)	23 ± 3.2 (22.3–23.8)	3.9 ± 3.7 (22.2–25.6)	22.8 ± 3.1 (22–23.6)	0.185
Daily walking device[b] (Cane)	10 (12.2)	7 (35)	3 (4.8)	< 0.001*

Notes: Participants who could lift the body up from the floor successfully in at least two over the three trials were arranged into the "pass" group, if not, they were placed in the "fail" group. [a]Data are presented as mean ± SD (95% confidence intervals), and compared between the pass and fail groups using the independent samples *t*-test. [b]The data are presented using number (%) and compared between the groups using the *Chi* square test. *Indicated significant differences between the groups.

Table 3. Data comparisons between participants who passed and failed a one-time seated push up test (1SPUT).

Variable		Total ($n = 82$)	Fail ($n = 20$)	Pass ($n = 62$)	P-value
Upper limb loading during a seated push up test (ULL-SPUT)					
	ULL-SPUT (BW%)	82.3 ± 9.8 (80.1–84.5)	71.2 ± 12.9 (65.2–77.2)	85.9 ± 5.1 (84.6–87.2)	< 0.001[a]
Standardmeasures					
Body composition[c]	Lean body mass (BW%)	65.1 ± 7.5 (63.1–67.1)	60.5 ± 5.1 (56.6–64.5)	66.0 ± 7.6 (63.8–68.2)	0.016[a]
	Bone mineral content (BW%)	3.4 ± 0.7 (3.3–3.6)	3.2 ± 0.3 (2.9–3.4)	3.5 ± 0.7(3.3–3.7)	0.044[a]
	Fat mass (BW%)	31.4 ± 8 (29.3–33.5)	36 ± 6 (31.4–40.7)	30.5 ± 8 (28.2–32.9)	0.033[a]
Muscle strength	Hand grip test (kg)	20.3 ± 5.4 (19.2–21.5)	16.4 ± 3.7 (14.6–18.1)	21.6 ± 5.3 (20.3–23)	< 0.001[a]
	Back extensor (kg)	26 ± 14.5 (22.8–29.2)	19.6 ± 10.6 (14.6–24.5)	28.1 ± 15 (24.3–31.9)	0.037[b]
	Leg extensor (kg)	30.1 ± 18 (26.1–34)	20 ± 12.3 (14.3–25.8)	33.3 ± 18.5 (28.6–38)	0.02[b]
Mobility	Five times sit-to-stand test (s)	13 ± 3.1 (12.4–13.7)	16.5 ± 3.2 (15–18)	11.9 ± 2 (11.4–12.4)	< 0.001[a]
	Timed up and go test (s)	13.9 ± 3.2 (13.2–14.6)	16.1 ± 3.2 (14.6–17.6)	13.2 ± 2.9 (12.5–14)	< 0.001[b]
	10-m walk test (m/s)	1 ± 0.2 (0.9–1)	0.8 ± 0.2 (0.7–0.9)	1 ± 0.2 (1–1.1)	< 0.001[a]
	2-min step test (times)	54.7 ± 14.8 (51.5–57.9)	47.1 ± 16.3 (39.5–54.7)	57.2 ± 13.5 (53.7 ± 60.6)	0.007[a]

Notes: The data are presented using mean ± standard deviation (95% confidence interval), BW = bodyweight, s = second, m = meter. Superscripts indicates the p-values from [a]the independent samples t-test, [b]the Mann-Whitney U-test, [c]there were 57 participants in this variable as it was additionally included after initiation of the study, with 48 participants passed and 9 participants failed a 1SPUT.

Table 4. The correlation between outcomes of seated push up tests and standard measures, including body compositions, muscle strength and mobility measures of the participants.

	Body compositions (g)			Muscle strength (kg)			Mobility			
Variable	Lean body mass	Bone mineral content	Body fat mass	Handgrip	Back muscles strength	Leg muscles strength	Five times sit-to-stand test (s)	Timed up and go test (s)	10-m walk test (m/s)	2-min step test (times)
ULL-SPUT[a](kg)	**0.785****	**0.628****	**0.515****	**0.547****	**0.456****	**0.345****	**−0.416****	**−0.288****	**0.332****	**0.247***
p-value	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**0.001**	**< 0.001**	**0.009**	**0.002**	**0.025**
5SPUT[b] (s)	−0.0281	**−0.306***	−0.059	−0.182	−0.115	−0.059	**−0.393****	0.173	**0.364****	**0.398****
p-value	0.053	**0.035**	0.691	0.156	0.372	0.650	**0.002**	0.178	**0.004**	**0.001**
10SPUT[b] (s)	−0.166	−0.164	−0.032	−0.049	−0.149	−0.126	**−0.355****	0.207	**0.361****	**0.357****
p-value	0.261	0.264	0.828	0.703	0.247	0.329	**0.005**	0.107	**0.004**	**0.004**
1minSPUT[b] (times)	**0.325***	0.244	−0.100	**0.356****	**0.425****	**0.466****	**−0.526****	**−0.306***	**0.332****	**0.421****
p-value	**0.024**	0.095	0.498	**0.005**	**0.001**	**< 0.001**	**< 0.001**	**0.016**	**0.008**	**0.001**

Notes: p-values were derived from Pearson correlation coefficients. Bold characters indicate the data with significant correlation at $^{*}p < 0.05$ and $^{**}p < 0.01$. [a]This variable was analysed in 57 participants for body compositions, and 82 participants for muscle strength and mobility measures. [b]These variables were analysed in 48 participants for body compositions, and 62 participants for muscle strength and mobility measures. ULL-SPUT: Upper limb loading during a seated push up test; 5SPUT: five-time seated push-up test; 10SPUT: 10-time seated push-up test; 1minSPUT: 1-min seated push-up test.

The body composition, muscle strength, and mobility also showed significant differences between the groups ($p < 0.05$; Table 3). The findings further indicated a moderate to strong correlation between the ULL-SPUT data, and all body composition ($r = 0.515$–0.785; $p < 0.001$; Table 4), muscle strength, and mobility measures ($r = 0.247$–0.547; $p < 0.05$) of the participants (Table 4).

All participants who passed the 1SPUT could complete other forms of SPUTs ($n = 62$). Their average 5SPUT time was 9.2 ± 4.1 s (95% confidence interval [CI]: 8.6–9.7 s), 10SPUT time was 17.7 ± 3.5 s (95% CI: 16.9–18.7 s), and 1minSPUT repetitions was 28.3 ± 8.0 times (95% CI: 26.4–30.3 times). There were no adverse events related to the SPUTs reported by any of the participants. The outcomes of the 1minSPUT showed a significant low to moderate correlation to LBM and muscle strength tests and mobility measures ($r = -0.306$–0.526; $p < 0.05$; Table 4) but not to BMC or body fat mass ($p > 0.05$). By contrast, the 5SPUT and 10SPUT data showed a significant low correlation to only the five times sit-to-stand test (FTSST), 10-m walk test (10MWT) and 2-min step test (2MST) ($r = -0.355$–0.398; $p < 0.01$; Table 4).

Discussion

This study assessed the validity and feasibility for various types of SPUTs among community-dwelling older people. All participants could complete the 1SPUT and ULL-SPUT, wherein a pass or fail in a 1SPUT could clearly discriminate participants with different body composition, muscle strength, and mobility outcomes ($p < 0.05$; Table 3). In addition, the ULL-SPUT showed a significant correlation to all body composition, muscle strength, and mobility measures (Table 4). Among the time-based SPUTs, the 1minSPUT showed significant correlation with LBM and all muscle strength and mobility measures, whilst the 5SPUT and 10SPUT showed a significant low correlation with only some mobility measures (Table 4).

Of all forms of SPUTs investigated in this study, the 1SPUT along with its ULL-SPUT are the least demanding measures and, thus they could be completed by all participants with poor and good functional ability (Table 3). Participants who passed a 1SPUT had the ULL-SPUT approximately 85% of their bodyweight which was significantly greater than that of those who failed the test (at approximately 71% of their bodyweight; $p < 0.001$; Table 3). Previous studies reported that the ability to increase the ULL-SPUT requires the complex interaction of many upper limb and upper trunk muscles, as well as perceptual information working cooperatively to generate muscle force and joint torque to lift the body upward by both arms.[9–11] Such ability requires SMM, a major part of LBM, in order for the muscles involved in the task to convert chemical energy into mechanical energy for force and power generation.[10,19] Muscular contraction also imposes mechanical loading onto the bones, along with cardiovascular stress, while body fat mass acts as resistance when completing the task.[19,20] Therefore, the 1SPUT and ULL-SPUT outcomes significantly correlated to body composition of the arms. Then the age-related physiological changes occurring throughout the body and in all body systems enabled outcomes of the tests involving upper limb and upper trunk muscles to reflect standard measures involving other body parts.[1,2] Consequently, the present findings indicate a significant correlation between the ULL-SPUT and all body composition, muscle strength, and mobility measures investigated in this study (Table 4). Furthermore, the participants who passed the 1SPUT had the outcomes of body composition, muscle strength, and mobility measures significantly better than those of the participants who failed the test ($p < 0.05$; Table 3).

Nonetheless, outcomes of the 1SPUT and ULL-SPUT may face ceiling effects in older individuals with good ability (i.e., score limitation at the top of a scale, i.e., always getting a *pass* in the test or nearly 100% of their bodyweight[35]). Thus the 1SPUT and ULL-SPUT outcomes in these cases may not represent the changes occurring in these participants, even there is actual change in the participants' body composition, muscle strength, or mobility. In such cases, other timed-based SPUTs (e.g., the 5SPUT, 10SPUT, and 1minSPUT)—which can be applied only in those who pass the 1SPUT—may be utilised to further challenge the ability of these individuals. However, the 5SPUT and 10SPUT, which can be completed within a short duration, may be unable to clearly reflect the variability in ability level among these participants. Therefore, the 5SPUT and 10SPUT outcomes showed a low correlation to only some mobility measures, including the FTSST, 10MWT, and 2MST ($r = -0.355$–0.398; $p < 0.01$; Table 4). On the contrary, outcomes of the most challenging

form of SPUT investigated in this study—namely, the 1minSPUT—might be able to detect the variability of participants with good functional ability. Therefore, 1minSPUT outcomes showed significant correlation with LBM, muscle strength and mobility of the participants ($r = -0.306$–0.526; $p < 0.05$; Table 4).

Previous studies have suggested that the slightest physical reduction could transform a person from independence into one with disability. Therefore, it is highly suggested to detect and monitor any abnormality early on, as well as to improve their physical conditions since they are still functioning independently, so that disability can be prevented or delayed.[36,37] The present findings suggest the feasibility and validity of the SPUTs—in various forms—as another practical upper limb measure to indicate body composition, muscle strength, and mobility, which are necessary for the independence and safety of community-dwelling older individuals. Such measures can be completed in a small area over a hard and smooth surface (e.g., over a bed); and thus they can be applied in various clinical, community-, and home-based settings.

However, there are some noteworthy limitations to this study. First, the study was conducted cross-sectionally among mostly well-functioning older participants who had a BMI of less than 30 kg/m^2. Thus, the findings may not clearly indicate the ability of the SPUTs to monitor changes in the body composition, muscle strength, and mobility of older individuals over time or of those who are frail or obese. Second, the SPUT measurements were taken using clinical push-up boards of a standard size, which may influence the outcomes of the test for participants with different heights. Third, only 57 participants (69.5%) were assessed for their body composition because this variable was additionally included after the initiation of the study. The lower number of participants as compared to that required in this study may affect clinical contribution for the findings of this variable, i.e., the correlation between SPUTs and body composition. Thus, a further study addresses all these limitations, as well as other psychometric properties needed for clinical application are still required to confirm the present findings and extend clinical benefit of SPUTs.

In conclusion, particular forms of SPUTs are feasible and valid measures to reflect body composition, muscle strength, and mobility among well-functioning older individuals. The 1SPUT and ULL-SPUT can be applied among those with good or poor functional ability, and they may be completed using a digital bathroom scale placed on a hard and even surface. The 1minSPUT—which may be assessed using an armchair or using small wooden boxes placed on a firm surface—can be employed to further challenge older individuals who have good functional ability. Such measures may be applied to early detect the abnormality relating to their muscle strength, mobility and body composition change that may occurred in older individuals, especially among those with lower limb limitations or in settings with limited area and equipment.

Conflicts of Interest

The authors have no conflicts of interest relevant to this paper.

Funding/Support

This study was supported by funding from the Royal Golden Jubilee (RGJ) Ph.D. programme, Thailand Research Fund (TRF, grant no. PHD/0174/2560). For the remaining authors none were declared.

Author Contributions

All authors were involved in concept, design, and planning of the study. In addition, all of them also took part in critical revision of the paper for intellectual content and finalised the manuscript. PP, PC and RI additionally contributed in collection and assembling of the data. PP and SP also took part in drafting of the manuscript. SA, SP and PA provided the study materials and technical support. SA additionally provided the funding support and administration of this project.

References

1. Amarya S, Singh K, Sabharwal M. Changes during aging and their association with malnutrition. J Clin Gerontol Geriatr 2015;6:78–84.
2. Colon- Emeric CS, Whitson HE, Pavon J, Hoenig H. Functional decline in older adults. Am Fam Physician 2013;88:388–94.
3. Frontera WR, Hughes VA, Fielding RA, Fiatarone MA, Evans WJ, Roubenoff R. Aging of skeletal

muscle: A 12-yr longitudinal study. J Appl Physiol 2000;88:1321–26.
4. Herman S, Kiely DK, Leveille S, O'Neill E, Cyberey S, Bean JF. Upper and lower limb muscle power relationships in mobility-limited older adults. J Gerontol A Biol Sci Med Sci 2005;60:476–80.
5. Puthoff ML, Nielsen DH. Relationships among impairments in lower-extremity strength and power, functional limitations, and disability in older adults. Phys Ther 2007;87:1334–47.
6. Shinohara M, Latash ML, Zatsiorsky VM. Age effects on force produced by intrinsic and extrinsic hand muscles and finger interaction during MVC tasks. J Appl Physiol 2003;95:1361–69.
7. Kwon I, Kim J-S, Shin C-H, Park Y, Kim J-H. Associations between skeletal muscle mass, grip strength, and physical and cognitive functions in elderly women: effect of exercise with resistive theraband. J Exerc Nutrition Biochem 2019; 23:50–5.
8. Garcia-Pena C, Garcia-Fabela LC, Gutierrez-Robledo LM, Garcia-Gonzalez JJ, Arango-Lopera VE, Perez-Zepeda MU. Handgrip strength predicts functional decline at discharge in hospitalized male elderly: a hospital cohort study. PLoS One 2013;8: e69849.
9. Short FX, Winnick JP. Test items and standards related to muscle strength and endurance on the Brockport Physical Fitness Test. Adapt Phys Activ Q 2005;22:371–400.
10. Yoshimura Y, Ise M. Analysis of the push-up movement base on action potentials of upper extremity muscles and ground reaction force between the palms and a force plate. Kawasaki J Med Welf 2005;11:1–7.
11. Wiyanad A, Amatachaya P, Sooknuan T, et al. The use of simple muscle strength tests to reflect body compositions among individuals with spinal cord injury. Spinal Cord 2021;60:99–105.
12. Poncumhak P, Wiyanad A, Siriyakul C, Kosura N, Amatachaya P, Amatachaya S. Validity and feasibility of a seated push-up test to indicate skeletal muscle mass in well-functioning older adults, Physiother Theory Pract 2022;1–8.doi:https://doi.org/10.1080/09593985.2021.2023931.
13. Dehghan M, Merchant AT. Is bioelectrical impedance accurate for use in large epidemiological studies? Nutr J 2008;9:26.
14. Mohamad AB, Nurakmal B. Sample size guideline for correlation analysis. World J Soc Sci Res 2016;3:37–46.
15. Whitney SL, Wrisley DM, Marchetti GF, Gee MA, Redfern MS, Furman JM. Clinical measurement of sit-to-stand performance in people with balance disorders: validity of data for the five-times-sit-to-stand test. Phys Ther 2005;85:1034–45.
16. Choi YJ. Dual-energy X-ray absorptiometry: Beyond bone mineral density determination. Endocrinol Metab (Seoul) 2016;31:25–30.
17. Bakkum AJ, Janssen TW, Rolf MP, et al. A reliable method for measuring proximal tibia and distal femur bone mineral density using dual-energy X-ray absorptiometry. Med Eng Phys 2014;36:387–90.
18. Blake GM, Fogelman I. The role of DXA bone density scans in the diagnosis and treatment of osteoporosis. Postgrad Med 2007;83:509–17.
19. Bohannon RW, Schaubert KL. Test-retest reliability of grip-strength measures obtained over a 12-week interval from community-dwelling elders. J Hand Ther 2005;18:426–27.
20. Roberts HC, Denison HJ, Martin HJ, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. Age Ageing 2011;40:423–29.
21. Bethards S, Everitt-Smith S, Roberts H, Scarborough G, Tate S, Bandy WD. Intrarater test-retest reliability of an instrument used to measure back and leg strength. Isokinet Exerc Sci 1995;5:31–5.
22. Ten Hoora GA, Muscha K, Meijerb K, Plasquia G. Test-retest reproducibility and validity of the back-leg-chest strength measurements. Isokinet Exerc Sci 2016;24:209–16.
23. Bohannon RW. Test-retest reliability of the five-repetition sit-to-stand test: a systematic review of the literature involving adults. J Strength Cond Res 2011;25:3205–7.
24. Goldberg A, Chavis M, Watkins J, Wilson T. The five-times-sit-to-stand test: validity, reliability and detectable change in older females. Aging Clin Exp Res 2012;24:339–44.
25. Thaweewannakij T, Wilaichit S, Chuchot R, et al. Reference values of physical performance in Thai elderly people who are functioning well and dwelling in the community. Phys Ther 2013;93:1312–20.
26. Podsiadlo D, Richardson S. The timed "Up & Go": A test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39:142–48.
27. Shumway-Cook A, Brauer S, Woollacott M. Predicting the probability for falls in community-dwelling older adults using the timed up & go test. Phys Ther 2000;80:896–903.
28. Steffen TM, Hacker TA, Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: Six-minute walk test, Berg balance scale, timed Up & go test, and gait speeds. Phys Ther 2002;82:128–37.
29. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. J Am Med Assoc 2011;305:50–58.
30. Lang JT, Kassan TO, Devaney LL, Colon-Semenza C, Joseph MF. Test-retest reliability and minimal

detectable change for the 10-meter walk test in older adults with Parkinson's disease. J Geriatr Phys Ther 2016;39:165–70.
31. Amatachaya S, Kwanmongkolthong M, Thongjumroon A, et al. Influence of timing protocols and distance covered on the outcomes of the 10-meter walk test. Physiother Theory Pract 2020;36:1348–53.
32. Middleton A, Fritz SL, Lusardi M. Walking speed: the functional vital sign. J Aging Phys Act 2015;23:314–22.
33. Haas F, Sweeney G, Pierre A, Plusch T, Whiteson J. Validation of a 2 minute step test for assessing functional improvement. Open J Ther Rehabil 2017;5:71–81.
34. Rikli RE, Jones CJ. Development and validation of a functional fitness test for community-residing older adults. J Aging Phys Act 1999;7:129–61.
35. Wang L, Zhang Z, McArdle JJ, Salthouse TA. Investigating ceiling effects in longitudinal data analysis. Multivariate Behav Res 2009;43(3):476–496.
36. Batsis JA, Daniel K, Eckstrom E, et al. Promoting healthy aging during COVID-19. J Am Geriatr Soc 2021;69(3):572–580.
37. Netz Y, Ayalon M, Dunsky A, Alexander N. The multiple-sit-to-stand' field test for older adults: what does it measure? Gerontology 2004;50:121–126.

Research Paper

Hong Kong Physiotherapy Journal
Vol. 42, No. 2 (2022) 137–149
DOI: 10.1142/S1013702522500135

Hong Kong Physiotherapy Journal
https://www.worldscientific.com/worldscinet/hkpj

Chinese (Mandarin) translation of the incremental shuttle walk test and its validity and reliability: A cross-sectional study

Wei Qin Ang[1], Hong Ting Tan[1], Si Min Goh[1], Samantha W. Seng[1], Katherin S. Huang[1,2], Melissa Y. Chan[1] and Meredith T. Yeung[1,*]

[1]*Health and Social Sciences Cluster*
Singapore Institute of Technology, Singapore

[2]*Department of Physiotherapy*
Khoo Teck Puat Hospital, Singapore

Meredith.Yeung@SingaporeTech.edu.sg

Received 30 July 2021; Accepted 19 July 2022; Published 31 August 2022

Background/Purpose: To date, there are no published validated Chinese versions of the incremental shuttle walk test (ISWT) instructions despite its wide clinical applications. Translation of the Chinese ISWT instruction is done in an *ad-hoc* manner within the Chinese-speaking populations, affecting the test's reliability and validity since translation can differ significantly between individuals. This warrants the need for psychometric testing of such translation.
Objectives: To develop a Chinese (Mandarin) version of the ISWT instructions (ISWT-CHN) that is conceptually equivalent to the original English version (ISWT-ENG) and establish its reliability and validity.
Methods: Forward and backward translations from the ISWT-ENG were done to generate the ISWT-CHN. Face and content validity was determined during the translation process. Intra-rater and inter-rater reliability of the ISWT-CHN, construct and criterion validity were established by analysing the ISWT and the gold standard cardiopulmonary exercise test results.
Results: The Item-Content validity index (I-CVI), Scale-level-Content validity index (S-CVI), and content validity ratio (CVR) of the ISWT-CHN were 1.0. Intra-class Correlation Coefficient (ICC) for inter-rater reliability between two raters were excellent (ICC = 0.99, 95% CI 0.97–1.0, $p < 0.001$; SEM = 0.85 m, MDC = 2.35 m). The intra-rater reliability of both Raters A (ICC = 0.92, 95% CI 0.53–0.98, $p = 0.003$;

*Corresponding author.

Copyright@2022, Hong Kong Physiotherapy Association. This is an Open Access article published by World Scientific Publishing Company. It is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) License which permits use, distribution and reproduction, provided that the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

SEM = 35 m, MDC = 97 m) and B (ICC = 0.90, 95% CI 0.76–0.96, $p < 0.001$; SEM = 32 m, MDC = 88 m) were good. In a sample of 32 healthy participants, both ISWT-CHN and ISWT-ENG instruction results showed low-positive correlations with the VO_{2max} determined from the cardiopulmonary exercise test ($r = 0.439$, $p < 0.001$; $r = 0.448$, $p < 0.001$). There is a very high correlation between ISWT-ENG and ISWT-CHN results with no statistically significant differences ($r = 0.967$, $p < 0.001$). The construct and criterion validity of the ISWT-CHN were established.

Conclusion: This study developed the ISWT-CHN and showed that it is a valid and reliable measure conceptually comparable to the ISWT-ENG. It will benefit the determination of functional exercise capacity in Chinese-speaking populations.

Key messages

- This study is aimed to develop a Chinese (Mandarin) version of the ISWT instructions.
- The ISWT Chinese translation is valid and reliable that is conceptually comparable to the original English instruction.
- The translated ISWT-Chinese instruction will enable the use of ISWT among the Chinese-speaking populations.

Keywords: Chinese translations; cross-sectional studies; incremental shuttle walk test; reproducibility of results.

Introduction

The Incremental Shuttle Walk Test (ISWT) is a commonly-used maximal field test that measures the total distance covered by the number of completed shuttles (ISWD).[1] The ISWT was first used to determine the functional exercise capacity in people with Chronic Obstructive Pulmonary Disease (COPD),[1] and was proven to be reliable and valid.[2–5] The application of ISWT was then expanded and proven to be reliable and valid in populations with other health conditions such as cardiovascular diseases,[6,7] lung cancer,[8] and peripheral arterial diseases.[9] Functional exercise capacity refers to the maximum amount of aerobic work an individual can sustain, defined by the maximal oxygen uptake ($VO_{2\,max}$). This variable is highly associated with a patient's ability to perform activities of daily living and the results are often used for exercise prescription among physiotherapists.[10] The ISWT is used to assess physical fitness and design personalised walking programmes,[11–13] as well as assess treatment outcomes and predict rehospitalisation, morbidity, and mortality rates.[14,15]

Traditionally, Cardiopulmonary Exercise Tests (CPETs) are the gold standard for measuring an individual's exercise capacity.[16] However, they require sophisticated equipment and trained staff to conduct the test and analyse its results. In contrast, the ISWT is simple and does not require specialised equipment or extensive training to conduct the tests. The externally-paced and incremental characteristics of the ISWT are also similar to laboratory CPETs,[1] making the ISWT an excellent alternative to assess exercise capacity in clinical settings.

The European Respiratory Society (ERS) and the American Thoracic Society (ATS) have established standardised English instructions for the ISWT.[17] However, its use with non-English speaking populations in regions and countries such as Hong Kong, Taiwan, Malaysia, China, and Singapore may be limited, given the difficulty in understanding the instructions. The Chinese (Mandarin) language is the second-most spoken language worldwide,[18] and approximately 75% of Singaporeans are Chinese, with 48% of them have Chinese as their primary language.[19] This highlights the need for Chinese instructions. However, there is no published validated Chinese version of the ISWT instructions to date. As a result, clinicians have to translate the ISWT instructions on an *ad-hoc* basis, which affects the reliability and validity of the test since translation can differ significantly between individuals, warranting the need for additional psychometric testing on an established measurement after cross-cultural translation and adaptation.[20] This is further attested by the results from the European Social Survey, which showed differences in measurement

quality across countries. Their study attempted to explain such differences and found that mistranslations and/or keywords lost in translations could affect the investigation results.[21] Therefore, this study aimed to develop a Chinese (Mandarin) version of the ISWT instructions (ISWT-CHN) that is conceptually equivalent to the original English version (ISWT-ENG) and subsequently established its reliability and validity.

Methods

Design

This was a translation and cross-sectional validity study assessing the reliability and validity of the ISWT after the adaptation to the Chinese (Mandarin) language. This study took place between June 2020 and April 2021. Ethical approval was obtained from the Institutional Review Board of the university (Number: 2020022), and all subjects provided written informed consent before participating in the study.

Development of ISWT-CHN

Before conceptualising this study, permission was first sought from the test originator.[1] The development of the ISWT-CHN adopted the cross-cultural adaptation process recommended by the World Health Organisation (WHO).[22] Figure 1 describes the translation and cross-cultural adaptation process of the ISWT-CHN. Two bilingual translators performed the forward translations independently. These two versions (Versions A and B) were combined to form the interim instructions (Version C), and it was then compared and revised by two investigators and the two forward translators in agreement to form the revised translated version (Version D). Twenty volunteers who were bilingual in English and Chinese but had no prior experience with the ISWT performed the backward translation of the revised-Chinese version (Version D) to English (Versions E1–E20). Lastly, the panel of researchers, which consisted of four investigators and two translators, reviewed the 20 backward translated versions for face and content validity and modified them to produce the final ISWT-CHN instructions (Version F).

Participants, investigators and centres

The participants were recruited via convenience sampling, with the following inclusion criteria: (1) Between 21 and 65 years old; (2) ambulant and not

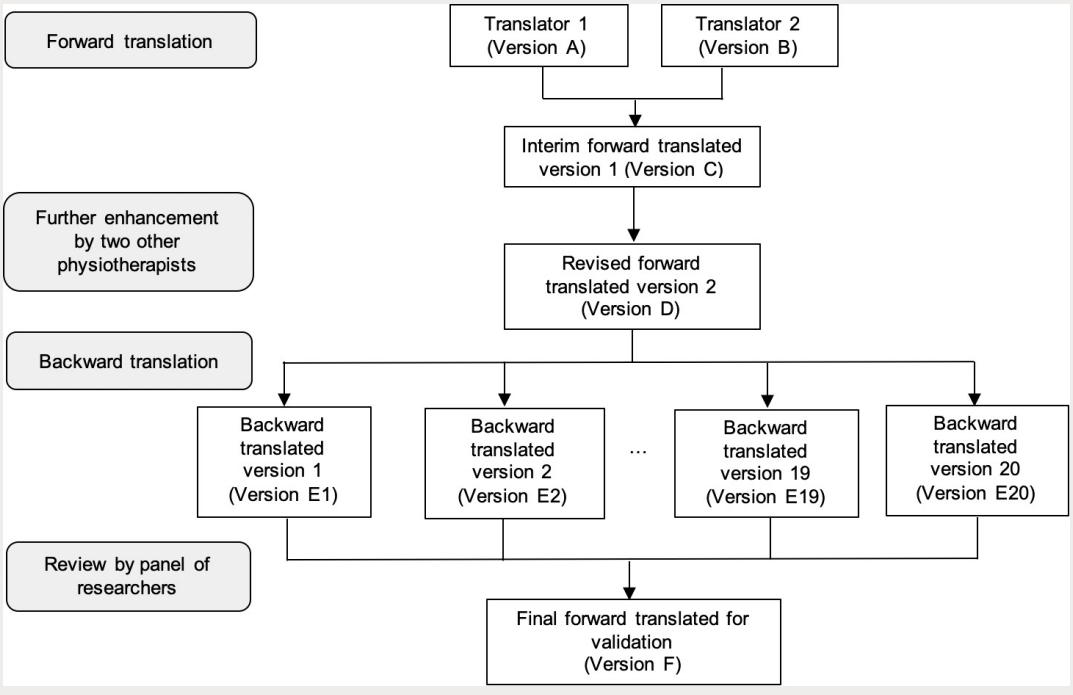


Fig. 1. Flow diagram of the translation and cross-cultural adaptation process.

using any walking aids; (3) colloquially proficient in Chinese and English; (4) has adequate mental capacity and able to follow instructions; and (5) was not involved with the backward translation in the development of the ISWT-CHN. Participants were excluded from the study if they had any cardiovascular, respiratory, neuromuscular and/or musculoskeletal disorder(s) that could hinder their ability to exercise. The Physical Activity Readiness Questionnaire for Everyone (PAR-Q+)[23] was used for screening. Prospective participants were excluded if they were presented with any contraindications for CPET[24] or spirometry results that suggested any potential airflow limitation, defined as a forced expiratory volume in one second (FEV_1) to forced vital capacity (FVC) ratio of less than 0.70.[25,26] According to other cross-cultural adaptation studies[27–31] and the minimum sample size suggested by statisticians,[32] 30 participants would provide adequate study power to observe the validity and reliability. All exercise testing, namely ISWT and CPET, were conducted by two investigators (Rater A and Rater B), undergraduate physiotherapy students with proficiency in conducting the tests, under close supervision from the principal investigator. All testing procedures and data collection took place in the exercise laboratory of the university.

The ISWT protocol

With adherence to the Technical Standard,[17] the ISWT was conducted with an open 10 m course marked by two cones that were placed 0.5 m inwards from either end. Participants were required to listen to a pre-recorded standardised ISWT instruction in either English or Chinese (Mandarin) before the test started and had to keep up with the walking speed dictated by the pre-recorded audio signals while walking up and down the course during the test. The test was terminated if the participant was limited by (1) fatigue; (2) dyspnoea; or (3) inability to maintain the required speed and failed to complete two consecutive shuttles. The distance covered was calculated from the total number of completed shuttles (ISWD). Parameters such as oxygen saturation (SpO_2), heart rate (HR), blood pressure (BP), and the modified Borg's scale (0–10) for Dyspnoea and Rate of Perceived Exertion (RPE)[33,34] were taken at various time points of the test. There was a minimum 30 min break between each test, and the SpO_2, HR, and BP had to return to baseline before the subsequent trial.

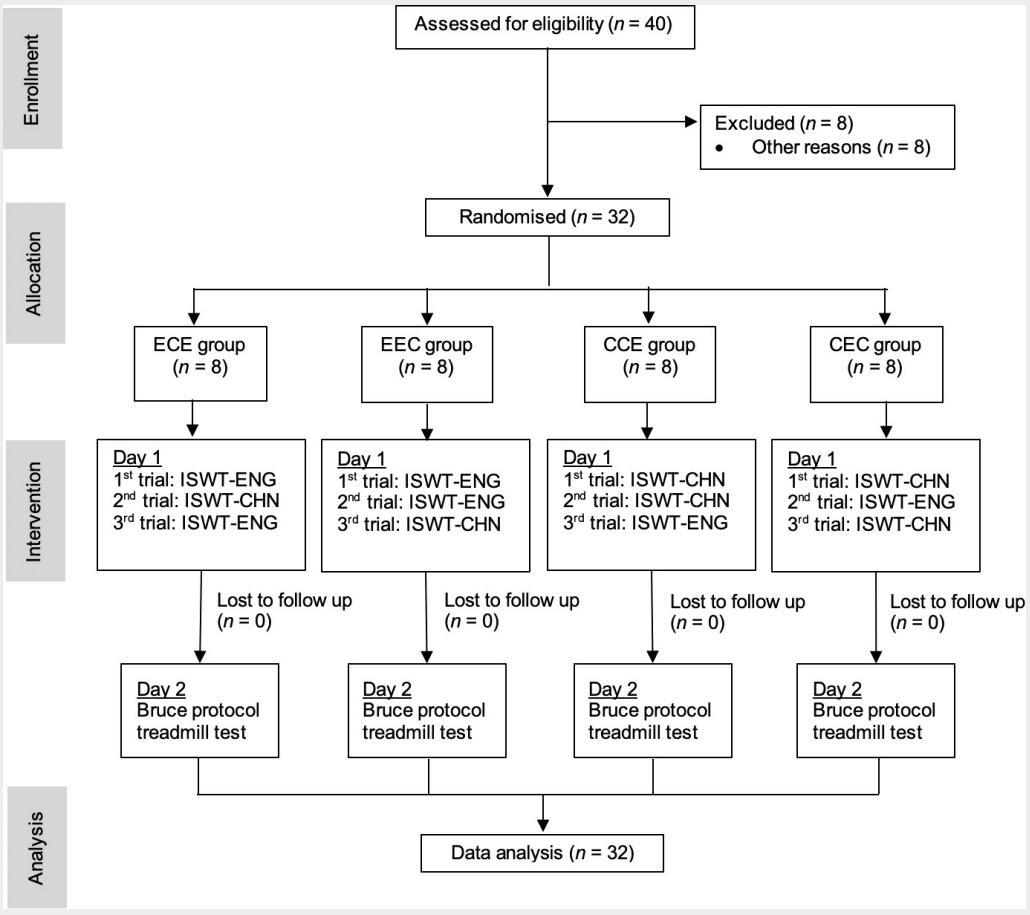
Bruce protocol treadmill test

The CPET was conducted on a treadmill (h/p/cosmos quasar® med, Germany) with metabolic gas analysis (COSMED Quark CPET Gas Analyser, Germany). The CPET was conducted using the standard Bruce treadmill protocol.[35] Participants were instructed to continue the test until maximal exhaustion while trained investigators monitored the necessary parameters [HR and rhythm via a 12-lead electrocardiogram, BP and RPE (0–10)]. The maximal oxygen consumption (VO_2max) recorded by the cardiopulmonary diagnostic software (OMNIA 1.6.3, COSMED, Rome, Italy) was used for data analysis. The CPET results served as the benchmark for criterion validity testing between ISWT-ENG and ISWT-CHN. The ISWT and CPET testings' were conducted on separate days.

Reliability and validity of ISWT-CHN

A pilot trial was first conducted on six healthy participants who fulfilled the same aforementioned inclusion and exclusion criteria with the ISWT-CHN instructions (Version F) to finalise the translation and establish the inter-reliability of the ISWT-CHN (Raters A and B). Both investigators contributed to the collection of the demographic data. Each participant completed a total of five trials on the same day. Considering the learning effect on the ISWT,[36,37] a practice trial was incorporated before the participants were randomised into A-B-B-A ($n = 3$) or B-A-A-B ($n = 3$) sequence via a computer-generated list for optimal efficiency and inter-rater reliability estimates.[38] The better of the two ISWD from each rater was used to establish the inter-rater reliability of the ISWT-CHN. No further changes to the ISWT-CHN instructions were made after the pilot trial, and Version F was concluded as the final ISWT-CHN instructions. Subsequently, the intra-rater reliability was established with an additional 14 healthy participants.

In the validity study, participants performed three trials of ISWT on the same day and the CPET on the following day. The block randomisation method with a block size of three was used. The block size of three was the minimum



Notes: ECE = English–Chinese–English, EEC = English–English–Chinese, CEC = Chinese–English–Chinese, CCE = Chinese–Chinese–English, ISWT-CHN = Chinese (Mandarin) version of the Incremental Shuttle Walk Test, ISWT-ENG = original version of the Incremental Shuttle Walk Test.

Fig. 2. Design and flow of participants through the validity study.

permutation for two variables, yet it was able to reduce bias and achieve balance in the allocation of participants compared to simple randomisation, particularly applicable with a small sample size.[39] The recruited participants were randomised using a computer-generated list into four groups with different test sequences for the validity testing of the ISWT-CHN. The groups were as follows: The English–Chinese–English (ECE; $n = 8$), English–English–Chinese (EEC; $n = 8$), Chinese–English–Chinese (CEC; $n = 8$) or Chinese–Chinese–English (CCE; $n = 8$) groups (Fig. 2). The better result of the two trials in the same language and the result of the remaining trial in the other language within the same test sequence, were chosen for analysis to account for the learning effect.[36,37] For example, the better result of the two English ISWD and the Chinese ISWD from the ECE group were used for subsequent analysis.

Data analysis

Statistical analysis was performed with IBM SPSS® Statistics for Windows Version 26.0 (IBM Corporation, Armonk, New York, USA). The level of significance was set at $p < 0.05$. Demographic and anthropometric data of participants were examined for normal distribution using the Shapiro–Wilk test. ANOVA was used to analyse normally distributed continuous variables between groups (Height, Weight, FEV_1/FVC ratio, Chinese and English ISWD, Absolute $VO_{2\,max}$) and the Kruskal–Wallis test was used to analyse non-normal continuous variables (Age). The Fisher's exact test was used to analyse sex differences between groups.

The inter-rater and intra-rater reliability were analysed via the interclass correlation (ICC) with a two-way mixed-effects model and a Bland–Altman plot.[40] The ICC was based on a 95% Confidence Interval (95% CI). Absolute reliability

was calculated with the standard error of measurement (SEM), minimal detectable change (MDC), and 95% limits of agreement (LOA). SEM was calculated based on the formula SD $\sqrt{(1 - \text{ICC})}$, while MDC at 95% level of confidence was calculated based on the formula $\text{SEM} \times \sqrt{2} \times 1.96$. The 95% LOA provides an interval within which 95% of differences between the two measurements are expected to lie. The alternate form of reliability between the two measurements can be ensured if the 95% LOA is narrow and at least 95% of the points lie within the limits.[41,42]

Face validity was determined during the translation process. Content validity was determined by the CVI and CVR.

(1) CVI is the most widely reported method for determining content validity in instrument development that examines its relevance and clarity, and it can be calculated using the Item-CVI (I-CVI) and the Scale-level-CVI (S-CVI).[43] This study used a 4-point Likert scale (not at all, needs some revision, needs minor revision, complete), where the I-CVI was calculated by the total ratings scored by all the panel members ($n = 6$; four investigators and two translators) and divided by the total number of panel members. Where I-CVI is greater than 0.79, the item is relevant; between 0.70 and 0.79, the item requires revisions; and when it is less than 0.70, the item is eliminated.[43,44] Similarly, S-CVI is determined by the number of items in the instrument that received a "highly complete" grade. The Universal Agreement (UA) among the panel members (S-CVI/UA) and the Average CVI (S-CVI/Ave) are two ways of determining S-CVI, the latter being a less conservative method.[43] S-CVI/UA is calculated by the sum of all items with I-CVI equal to 1 divided by the total number of items, whereas S-CVI/Ave is equal to the sum of all the I-CVIs divided by the number of items. Content validity is excellent when the S-CVI/UA is more than 0.8 and the S-CVI/Ave is more than 0.9.[44] This study used the S-CVI/UA method.

(2) CVR measures the essentiality of an item[45] CVR ranges from −1 to 1, with a higher score indicating greater agreement among panel members. $\text{CVR} = (\text{Ne} - \text{N}/2)/(\text{N}/2)$, where Ne is the number of panellists who rated an item as "essential" and N is the total number of panellists.[43] Each sentence's essentiality was examined on a 3-point Likert scale (not essential, useful but not essential, essential). For a panel size of six, $\text{CVR} = 1.0$ ($p = 0.05$)[46] is required to be statistically significant.

Construct and criterion validities were analysed using Pearson's correlation coefficient r. Construct validity was determined from the correlation between the distances obtained from ISWT-ENG and ISWT-CHN, while criterion validity was determined from the correlations between the ISWDs of the ISWT-CHN, ISWT-ENG and $\text{VO}_{2\,\text{max}}$ of the CPET.

Results

Flow of participants

Six participants were recruited during the pilot trial to establish the inter-rater reliability, and an additional 14 participants were included to establish the intra-rater reliability. All screened participants fulfilled the inclusion criteria, and none were excluded according to the exclusion criteria. Table 1 presents the overall characteristics of the participants in the reliability study. Subsequently, 40 participants were assessed for eligibility during the validity study, and eight participants were excluded according to the exclusion criteria. The remaining 32 participants, 20 males (63%) and 12 females (37%) satisfied the inclusion and exclusion criteria and participated in the validity study. The demographic and clinical data (age, sex, height, weight, spirometry readings, ISWD, and $\text{VO}_{2\,\text{max}}$)

Table 1. Demographics of participants in the reliability study ($n = 20$).

	Participants
Age (years)	26.7 (SD 5.1)
Sex	
Male	11
Female	9
Height (cm)	169 (SD 7.6)
Weight (kg)	67 (SD 14.9)
FEV_1/FVC ratio (%)	86 (SD 5.0)

Notes: SD: Standard deviation; cm: centimetres; kg: kilograms; FEV_1/FVC ratio (%): forced expiratory volume in one second (FEV_1) to forced vital capacity (FVC) ratio in percentage.

Table 2. Demographics and clinical data of participants in the validity study ($n = 32$).

	ECE ($n = 8$)	EEC ($n = 8$)	CEC ($n = 8$)	CCE ($n = 8$)	p-value
Age (years)	24 (SD 2.2)	27 (SD 3.7)	27 (SD 6.5)	24 (SD 3.1)	0.23
Sex					0.96
Male ($n = 20$)	6	5	5	4	
Female ($n = 12$)	2	3	3	4	
Height (cm)	168 (SD 5.4)	165 (SD 6.0)	168 (SD 8.1)	170 (SD 7.3)	0.53
Weight (kg)	62 (SD 7.3)	60 (SD 9.2)	72 (SD 16.0)	62 (SD 12.0)	0.18
FEV_1/FVC ratio (%)	90 (SD 3.3)	85 (SD 6.0)	85 (SD 3.0)	88 (SD 6.5)	0.15
Chinese ISWD (m)	699 (SD 155.0)	801 (SD 181.0)	738 (SD 150.0)	803 (SD 163.0)	0.52
English ISWD (m)	716 (SD 136.0)	801 (SD 174.0)	713 (SD 160.0)	806 (SD 157.0)	0.47
Absolute $VO_{2\,max}$ (mL/min/kg)	47 (SD 8.7)	40 (SD 8.1)	41 (SD 14.0)	47 (SD 16.0)	0.52

Notes: ECE = English–Chinese–English, EEC = English–English–Chinese, CEC = Chinese–English–Chinese, CCE = Chinese–Chinese–English, n = number of participants, ISWD = Incremental Shuttle Walk Distance, FEV_1 = forced expiratory volume in one second, FVC = forced vital capacity, and $VO_{2\,max}$ = maximal oxygen consumption. *Significance of p values < 0.05.

of the participants in the validity study are summarised in Table 2.

Reliability of ISWT-CHN

Table 3 illustrates the inter- and intra-rater reliability. The ISWT-CHN displayed excellent inter-rater reliability (ICC = 0.99, 95% CI 0.97–1.0, $p < 0.001$) between Raters A and B, with the SEM and MDC being 0.85 metres (m) and 2.35 m, respectively. Figure 3 shows the Bland–Altman plot for the inter-rater reliability, and the dots on the plot represents each of the six participants for the pilot study. From the plot, all the points were within the 95% LOA (+95% LOA = 37 m, −95% LOA = −37 m). Good intra-rater reliability (Rater A: ICC = 0.92, 95% CI 0.53–0.98, $p = 0.003$; Rater B: ICC = 0.90, 95% CI 0.76–0.96, $p < 0.001$) was observed. The SEM values for the intra-rater reliability testing were 35 m for Rater A and 32 m for Rater B. The MDC values for the intra-rater reliability testing were 97 m for Rater A and 88 m for Rater B. The SEM values for the reliability testing could be considered small as they were below 10% of the highest achievable distance of 1020 m.[47]

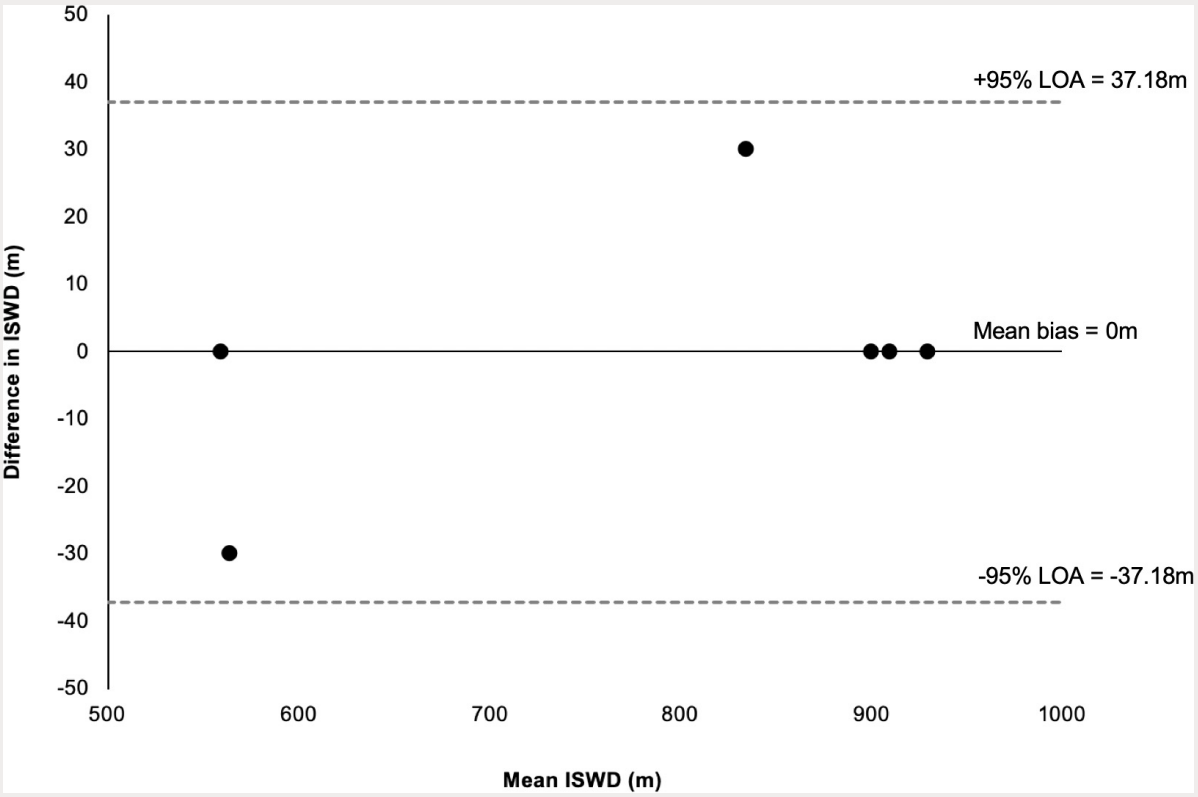
Validity of ISWT-CHN

The rigorous translation process established the face validity of the ISWT-CHN according to the translation and cross-cultural adaptation guidelines by the WHO.[22] The panel of six investigators synonymously agreed that the face meaning of the 20 copies of backward translations collected was equivalent to the original English instructions. Subsequently, content validity was established via quantitative and qualitative assessments.

Table 3. Reliability of the ISWT-CHN.

	ICC	95% CI	p-values	SEM (m)	MDC (m)
Inter-rater reliability ($n = 6$)	0.99	0.97-1.00	< 0.001*	0.85	2.35
Intra-rater reliability ($n = 20$)					
Rater A	0.92	0.53-0.98	$= 0.003$*	35.0	97.0
Rater B	0.90	0.76-0.96	< 0.001*	32.0	88.0

Notes: ISWT-CHN = Chinese (Mandarin) version of the Incremental Shuttle Walk Test, ICC = Intra-class Correlation Coefficient, CI = Confidence Interval, SEM = Standard Error of Measurement, and MDC = Minimal Detectable Change. *Significance of p values < 0.05.



Notes: The differences in incremental shuttle walk distances between raters are plotted against the mean scores. The straight line represents the mean difference between the two raters; dashed lines represent the 95% limits of agreement. ISWD = Incremental Shuttle Walk Distances and LOA = limits of agreement.

Fig. 3. Bland–Altman plot comparing the agreement between two raters.

Six investigators provided feedback and ratings on the essentiality, relevance, and clarity of each sentence of the translated instruction. There was unanimous agreement that the conceptual meaning of the IWST-CHN was consistent with the original version. The I-CVI, S-CVI/Ave and CVR of the ISWT-CHN were 1.0.

The ISWT-CHN had a mean distance of 760 m [standard deviation (SD) 161] and ISWT-ENG had a mean distance of 759 m (SD 156). The Pearson's coefficient of ISWT-CHN versus CPET ($r = 0.439$, $p < 0.001$) as compared to ISWT-ENG versus CPET ($r = 0.448$, $p < 0.001$). This similarly weak positive correlation with the $VO_{2\,max}$ established the criterion validity of the ISWT-CHN. Additionally, the ISWT-CHN had satisfactory construct validity, as shown by the very high positive correlation between the ISWD of the ISWT-CHN and ISWT-ENG ($r = 0.967$, $p < 0.001$). Figure 4 shows the scatterplots that depict the relationships between $VO_{2\,max}$ (mL/min/kg), ISWT-CHN distance (m) and ISWT-ENG distance (m).

Discussion

This study aimed to cross-culturally adapt the English ISWT instructions to Chinese (Mandarin) and evaluate the translation's reliability and validity. This study revealed that ISWT-CHN is a reliable and valid outcome measure. Intra-rater reliability of both Raters A and B were good, with an ICC of 0.92, $p = 0.003$ and 0.90, $p < 0.001$, respectively. Inter-rater reliability was excellent with an ICC of 0.99, $p < 0.001$, and the Bland–Altman plot revealed all points within ±1.96 SD of the mean difference implying that the two raters agreed. Absolute reliability measurements also showed that the ISWT-CHN results were precise and had low margins of error between raters, with an SEM of 0.85 m and an MDC of 2.35 m. This agrees with the results from other studies that showed good to excellent reliability, for instance, in the COPD population (ICC = 0.80–0.93)[2] and patients with peripheral obstructive arterial disease (ICC = 0.95).[9]

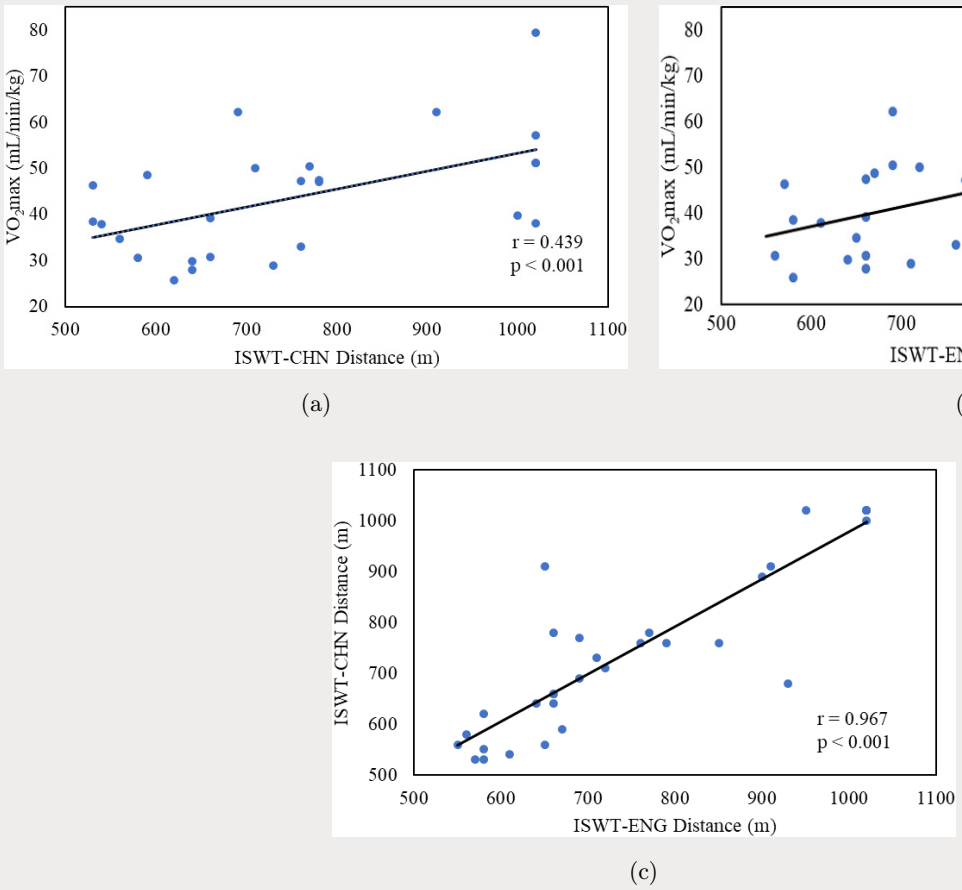


Fig. 4. Scatterplots of $VO_{2\,max}$, ISWT-CHN distance and ISWT-ENG distance. (a) $VO_{2\,max}$ and ISWT-CHN distance, (b) $VO_{2\,max}$ and ISWT-ENG distance, (c) ISWT-CHN distance and ISWT-ENG distance. ISWT-CHN = Chinese (Mandarin) version of the Incremental Shuttle Walk Test, ISWT-ENG = original version of the Incremental Shuttle Walk Test, $VO_{2\,max}$ = maximal oxygen consumption, r = Pearson's correlation, and p = probability value.

The validity of the ISWT-CHN was established. Face validity was determined during the translation process. Content validity was ascertained by the determination of CVI, in the form of I-CVI and S-CVI, as well as CVR. Construct and criterion validity of ISWT-CHN were also established. The construct validity of the ISWT-CHN was established through the excellent correlation with the ISWT-ENG ($r = 0.967$, $p < 0.001$). The ISWT-ENG and ISWT-CHN results were compared with CPET $VO_{2\,max}$ results, the gold standard to measure exercise capacity.[16] The criterion validity of the ISWT-CHN was indicated by the largely similar significant positive correlation of the ISWT-CHN and ISWT-ENG with the $VO_{2\,max}$($r = 0.439, p < 0.001; r = 0.448, p < 0.001$), which suggested strong agreement between both versions of ISWT. The similar degree of weak positive correlation between the $VO_{2\,max}$ and ISWT in both languages and the strong correlation between the ISWT-CHN and ISWT-ENG demonstrated the equivalence of the instructions in both the languages. Therefore, the establishment of criterion validity was inferred.

This study revealed a weaker correlation between ISWT results in both languages and CPET $VO_{2\,max}$ results ($r = 0.439$–0.448) as compared to other studies. The Pearson's correlation coefficient r ranged from 0.72 to 0.85 in the COPD population,[3–5] from 0.73 to 0.83 in patients with cardiac diseases,[6,7] and was 0.67 in lung cancer patients.[8] A possible reason for this is because the original ISWT[1] was developed for patients with COPD, and the protocol does not allow running to keep up with the set speeds. This, however, might not be sensitive to predict functional exercise capacity in the healthy population as they are less likely to reach 85% of the predicted maximal HR or become too breathless to continue just by walking. Moreover, when observing the best ISWT trial results

regardless of the language used in this study, none of the participants terminated the test due to excessive dyspnoea and/or leg fatigue.[1] In fact, most participants (78%, $n = 25$) terminated the test due to the inability to maintain the required speed.[1] Hence, in most participants, functional exercise capacity could not be accurately predicted as the test was terminated before adequate cardiovascular stress. The remaining 22% of participants ($n = 7$) completed all 12 levels at least once out of the three ISWT trials, and 86% ($n = 6$) of them were males. Functional exercise capacity could not be accurately predicted in this subgroup of participants due to ceiling effects. To overcome these limitations when using ISWT in a healthy population, Probst *et al.*[13] suggested allowing the participants to run if needed and increasing the number of stages beyond 12 levels by increasing the speed by 0.17 metres per second (m/s) each minute. However, this modified protocol[37] is not commonly used in clinical settings.

The ISWT is widely used in research and rehabilitation as a measure of exercise capacity. It is often chosen over other measures of exercise capacity, such as the 6-min walk test (6MWT), as this test necessitates at least 30 m of walking space[48] which may not be available at clinical or research facilities. Furthermore, given that the 6MWT is associated with several limitations related to its self-paced nature, the ISWT is superior when determining exercise intensity because of its progressive nature.[48,49] The ISWT is relevant and particularly useful when prescribing exercise intensity using the percentage of peak performance[48] and thus may be more beneficial in the rehabilitation or general health promotion setting when tailoring individualised exercise intervention.[11–13] The newly translated Chinese instructions of the ISWT may offer a reliable and validated method of field-based exercise testing, benefiting Chinese-speaking populations around the globe.

There are several limitations of this study. Firstly, the study used convenience sampling and recruited only a small sample size, which might have introduced bias where specific subgroups of populations might have been over- or under-represented. For instance, our participants were relatively young, with a mean age of 26.7 (SD 5.1) and 25 (SD 4.2) years in the reliability and validity studies, respectively, although the study's age inclusion criteria were set between 21 and 65 years old. This also resulted in a non-normal distribution with data skewed to the left. This may be due to the study's recruitment process, which required the participants to be on social media to obtain information about the study. This method may have limited the study's reach to only those proficient in technology navigating the social media platforms. Furthermore, a portion of interested participants who were in older age groups chose not to participate in the study after understanding what CPET was, and declined with the main reason being "afraid of over-exerting". Although interested and willing to perform CPET, some older adults were excluded from the study as they had chronic medical conditions and were deemed unsafe for exercise until certified by qualified personnel, as stated in the PAR-Q$^+$. Additionally, it is worth noting that the study was conducted amid the COVID-19 global pandemic, where multiple measures have been put in place, such as movement restrictions and proper social distancing to curb transmissions. Thus, participant recruitment was mainly limited to the younger population, and the sample size fell short of the 30 participants required to provide adequate study power. Notwithstanding, the reliability results are promising despite the small sample size. It is also unclear if there exists a difference in colloquial usage and understanding of the language between age groups which warrants further exploration.

Nonetheless, this study is the first to develop a Chinese version of the ISWT instructions with established reliability and validity. The study results were comparable to other published data, and it showed significant results to confirm that the ISWT-CHN was conceptually equivalent to its original English instructions. Thus, this study can serve as a framework for future studies with larger sample sizes and participants of varied backgrounds despite its limitations.

Conclusion

This study translated and cross-culturally adapted the original English instructions of the ISWT into the Chinese (Mandarin) language and subsequently established its reliability and validity. The ISWT-CHN was shown to be a reliable and valid measurement of exercise capacity with psychometric properties similar to the original English version.[1] This development is helpful to the Chinese-speaking population as clinicians can now

conduct the ISWT using standardised Chinese instructions instead of performing *ad-hoc* translation. However, given the limited sample size, future research on the ISWT-CHN with a larger sample size and participants of different backgrounds is recommended.

Acknowledgment

We are grateful to the participants who took their time off to participate in this study.

Conflict of Interest

The author(s) declared no potential conflicts of interest concerning this paper's research, authorship, and/or publication.

Funding/Support

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this paper. This work was supported by the Singapore Institute of Technology, Grant No. R-MOE-A404-H009.

Ethics Approval

The Singapore Institute of Technology-Institutional Review Board (SIT-IRB Project Number: 2020022) approved this study. All participants gave written informed consent before data collection began.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

Wei Qin Ang and Hong Ting Tan collected, analysed and interpreted the data, prepared and revised the paper critically for important intellectual content. Si Min Goh collected, analysed, and interpreted the data. Samantha W. Seng collected and interpreted the data. Katherin S. Huang and Melissa Y. Chan conceptualised and designed the study, drafted and revised the paper critically for important intellectual content. Meredith T. Yeung conceptualised and designed the study, analysed and interpreted the data, revised the paper critically for important intellectual content and approved the final paper. Wei Qin Ang and Hong Ting Tan are the co-first authors who contributed and credited equally to this work.

References

1. Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. Thorax 1992;47:1019–24.
2. Campo LA, Chilingaryan G, Berg K, Paradis B, Mazer B. Validity and reliability of the modified shuttle walk test in patients with chronic obstructive pulmonary disease. Arch Phys Med Rehabil 2006;87:918–22.
3. Onorati P, Antonucci R, Valli G, et al. Non-invasive evaluation of gas exchange during a shuttle walkingest vs. a 6-min walking test to assess exercise tolerance in COPD patients. Eur J Appl Physiol 2003;89:331–6.
4. Satake M, Shioya T, Takahashi H, Kawatani M. Ventilatory responses to six-minute walk test, incremental shuttle walking test, and cycle ergometer test in patients with chronic obstructive pulmonary disease. Biomed Res 2003;24: 309–16.
5. Turner SE, Eastwood PR, Cecins NM, Hillman DR, Jenkins SC. Physiologic responses to incremental and self-paced exercise in COPD: A comparison of three tests. Chest 2004;126:766–73.
6. Green DJ, Watts K, Rankin S, Wong P, O'Driscoll JG. A comparison of the shuttle and 6 minute walking tests with measured peak oxygen consumption in patients with heart failure. J Sci Med Sport 2001;4:292–300.
7. Lewis ME, Newall C, Townend JN, Hill SL, Bonser RS. Incremental shuttle walk test in the assessment of patients for heart transplantation. Heart 2001;86:183–7.
8. Win T, Jackson A, Groves AM, Sharples LD, Charman SC, Laroche CM. Comparison of shuttle walk with measured peak oxygen consumption in patients with operable lung cancer. Thorax 2006;61:57–60.
9. da Cunha-Filho IT, Pereira DA, de Carvalho AM, Campedeli L, Soares M, de Sousa Freitas J. The reliability of walking tests in people with claudication. Am J Phys Med Rehabil 2007;86:574–82.
10. Arena R, Myers J, Williams MA, et al. Assessment of functional capacity in clinical and research settings. Circulation 2007;116:329–43.

11. Harrison SL, Greening NJ, Houchen-Wolloff L, et al. Age-specific normal values for the incremental shuttle walk test in a healthy British population. J Cardiopulm Rehabil Prev 2013;33:309–13.
12. Itaki M, Kozu R, Tanaka K, et al. Reference equation for the incremental shuttle walk test in Japanese adults. Respir Investig 2018;56:497–02.
13. Probst VS, Hernandes NA, Teixeira DC, et al. Reference values for the incremental shuttle walking test. Respir Med 2012;106:243–8.
14. Emtner MI, Arnardottir HR, Hallin R, Lindberg E, Janson C. Walking distance is a predictor of exacerbations in patients with chronic obstructive pulmonary disease. Respir Med 2007;101:1037–40.
15. Nutt CL, Russell JC. Use of the pre-operative shuttle walk test to predict morbidity and mortality after elective major colorectal surgery. Anaesthesia 2012;67:839–49.
16. Ferrazza AM, Martolini D, Valli G, Palange P. Cardiopulmonary exercise testing in the functional and prognostic evaluation of patients with pulmonary diseases. Respiration 2009;77:3–17.
17. Holland AE, Spruit MA, Singh SJ. How to carry out a field walking test in chronic respiratory disease. Breathe (Sheffield, England) 2015; 11:128–39.
18. Statista. The most spoken languages worldwide in 2021. https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/: Statista, 2021.
19. DoS. Singapore Census of Population 2010, Statistical Release 1. Demographic Characteristics, Education, Language and Religion. Department of Statistics, 2011.
20. Peña ED. Lost in translation: Methodological considerations in cross-cultural research. Child Dev 2007;78:1255–64.
21. Oberski D, Saris WE and Hagenaars J. Why are there differences in measurement quality across countries. In: Loosveldt G, Swyngedouw M and Billiet J, (eds.). Measuring Meaningful Data in Social Research, 1st ed. Leuven: Acco, 2007.
22. WHO. Process of translation and adaptation of instruments. Wold Health Organization, 2020.
23. Warburton DE, Jamnik VK, Bredin SS, Gledhill N. The physical activity readiness questionnaire for everyone (PAR-Q+) and electronic physical activity readiness medical examination (ePARmed-X +). Health Fit J Canada 2011;4:3–17.
24. ATS/ACCP Statement on cardiopulmonary exercise testing. Am J Respir Crit Care Med 2003;167:211–77.
25. Rabe KF, Hurd S, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med 2007;176:532–55.
26. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. Eur Respir J 2005;26:319–38.
27. Macfarlane DJ, Lee CC, Ho EY, Chan KL, Chan DT. Reliability and validity of the Chinese version of IPAQ (short, last 7 days). J Sci Med Sport 2007;10: 45–51.
28. Poon DM, Chow LC, Au DK, Hui Y, Leung MC. Translation of the dizziness handicap inventory into Chinese, validation of it, and evaluation of the quality of life of patients with chronic dizziness. Ann Otol Rhinol Laryngol 2004;113:1006–11.
29. Dhippayom JP, Trevittaya P, Cheng ASK. Cross-cultural adaptation, validity, and reliability of the patient-rated Michigan hand outcomes questionnaire for Thai patients. Occup Ther Int 2018;2018:8319875.
30. Passos L, Tavares J, Figueiredo D. Cross-cultural adaptation and analysis of the psychometric properties of the Portuguese version of the Edinburgh feeding evaluation in Dementia scale. Port J Public Health 2020;38:142–50.
31. Leung RW, Leung M-L, Chung P-K. Validity and reliability of a Cantonese-translated rating of perceived exertion scale among Hong Kong adults. Percept Mot Ski 2004;98:725–35.
32. Onwuegbuzie AJ, Collins KM. A typology of mixed methods sampling designs in social science research. Qual Rep 2007;12:281–16.
33. Borg GA. Perceived exertion. Exerc Sport Sci Rev 1974;2:131–53.
34. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc. 1982;14:377–81.
35. Bruce RA, Blackmon JR, Jones JW, Strait G. Exercising testing in adult normal subjects and cardiac patients. Pediatrics 1963;32:742–56.
36. Dyer F, Marriner P, Cheema K, Bott J. Is a practice incremental shuttle walk test really necessary? Chronic Respir Dis 2011;8:201–5.
37. Bradley J, Howard J, Wallace E, Elborn S. Validity of a modified shuttle test in adult cystic fibrosis. Thorax 1999;54:437–9.
38. Hallgren KA. Computing inter-rater reliability for observational data: An overview and tutorial. Tutor Quant Meth Psychol. 2012;8:23–34.
39. Efird J. Blocked randomization with randomly selected block sizes. Int J Environ Res Public Health. 2010;8:15–20.
40. Bland JM, Altman DG. Measurement error. Bmj 1996;313:744.
41. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998;26:217–38.
42. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med 2000;30:1–15.

43. Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi-Majd H, Nikanfar A-R. Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. J Caring Sci 2015;4:165–78.
44. Polit DF, Beck CT. The content validity index: Are you sure you know what's being reported? critique and recommendations. Res Nursing Health 2006;29:489–97.
45. Yamada J, Stevens B, Sidani S, Watt-Watson J, De Silva N. Content validity of a process evaluation checklist to measure intervention implementation fidelity of the EPIC intervention. Worldviews Evid-Based Nurs 2010;7: 158–64.
46. Ayre C, Scally AJ. Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. Meas Eval Couns Dev 2013;47:79–86.
47. Prince B, Makrides L, Richman J. Research methodology and applied statistics. Part 2: The literature search. Physiother Can 1980;32:201–6.
48. Holland AE, Spruit MA, Troosters T, et al. An official European Respiratory Society/American Thoracic Society technical standard: Field walking tests in chronic respiratory disease. Eur Respir J 2014;44:1428–46.
49. Wise RA, Brown CD. Minimal clinically important differences in the six-minute walk test and the incremental shuttle walking test. J Chron Obstruct Pulmon Dis 2005;2:125–9.