

# Effectiveness, safety and patients' perceptions of an immersive virtual reality-based exercise system for poststroke upper limb motor rehabilitation: A proof-of-concept and feasibility randomized controlled trial

DIGITAL HEALTH  
Volume 9: 1–33  
© The Author(s) 2023  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/20552076231203599  
journals.sagepub.com/home/dhj



Jiayin Chen<sup>1</sup>, Calvin Kalun Or<sup>1</sup> , Zhixian Li<sup>2</sup>, Eric Hiu Kwong Yeung<sup>3</sup>, Yong Zhou<sup>2</sup> and Tiantian Hao<sup>2</sup>

## Abstract

**Objective:** This study aimed to examine the effectiveness, safety and patients' perceptions of an immersive virtual reality (VR)-based exercise system for poststroke upper limb rehabilitation.

**Methods:** A proof-of-concept, 2-week randomized controlled trial was conducted. Fifty stroke patients were randomly assigned to either use the immersive VR-based exercise system to perform upper limb exercises for 2 weeks (intervention) or play commercial games (control). Effectiveness, safety and patients' perceptions of the exercise system were assessed at baseline and at 1- and 2-week follow-ups.

**Results:** Intention-to-treat analysis revealed that after 2 weeks, statistically significant improvements in shoulder flexion active range of motion (AROM), shoulder abduction AROM, perceived upper limb motor function and quality of life (QoL) were observed in one or both groups, but not between the groups. Per-protocol analysis showed that after 2 weeks: (i) statistically significant improvement in shoulder abduction AROM was obtained in the intervention group, and the difference in the mean changes between the groups was statistically significant; (ii) statistically significant improvements in coordination/speed (Fugl-Meyer Assessment for Upper Extremity), shoulder flexion AROM, perceived upper limb motor function and QoL were obtained in one or both groups, but not between the groups.

**Conclusions:** The immersive VR-based exercise system is a potentially effective, safe and acceptable approach for supporting poststroke motor rehabilitation. These findings can serve as a basis for larger-scale studies on the application of VR for poststroke exercises.

## Keywords

Immersive virtual reality, stroke, upper limb motor exercises, rehabilitation

Submission date: 3 April 2023; Acceptance date: 7 September 2023

## Introduction

Upper limb motor impairment is a common disability after stroke. Afflicted patients need to undergo therapeutic exercises to restore motor function and independence. However, because such treatment is typically supervised by physical therapists and takes place in clinics, it can be costly and inconvenient to obtain.

Technologies such as virtual reality (VR) have been used to overcome these limitations, because VR can promote access to exercises, aid motor learning and enhance individuals' motivation to perform therapeutic exercises while also

<sup>1</sup>Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Hong Kong, China

<sup>2</sup>Department of Rehabilitation Medicine, Dingzhou People's Hospital, Dingzhou, China

<sup>3</sup>Department of Physiotherapy, The University of Hong Kong-Shenzhen Hospital, Shenzhen, China

### Corresponding author:

Calvin Kalun Or, The University of Hong Kong, Room 8-7, 8/F, Haking Wong Building, Hong Kong, China.  
Email: klor@hku.hk



providing a safe simulation of real-life daily activities, real-time multimodal feedback on performance and engaging exercising experience.<sup>1–9</sup> VR applications range from non-immersive to immersive, depending on the degree to which users are isolated from the real world when interacting with virtual environments. Non-immersive VR is achieved through a 2D screen, which allows users to interact with virtual environments from the ‘outside’ while having a full view of the surrounding real world, and interaction is typically limited to the use of a mouse or joystick.<sup>10–12</sup> Immersive VR is usually achieved through a head-mounted display (HMD), which isolates users from the actual world and immerses them inside stereoscopic environments where they can interact with the environments using their body.<sup>10,13</sup> Compared to non-immersive VR, immersive VR can incorporate multiple sensory and motor channels into experience, allow for more natural and intuitive interaction, provide a more realistic perception of space (e.g., depth, distances and size) and elicit a higher perceived level of presence.<sup>14–16</sup> Being immersed in virtual environments can arouse physiological responses,<sup>14,15</sup> activate brain activities,<sup>17</sup> and influence the effectiveness of VR-based treatments.<sup>18,19</sup>

Reviews of the use of VR in poststroke motor rehabilitation have shown its effectiveness in safely improving certain outcomes, such as upper limb motor function (as measured by the Fugl–Meyer Assessment for Upper Extremity [FMA-UE]) and independence in daily living (as measured by the Functional Independence Measure).<sup>20–24</sup> However, evidence on the effects of VR on other relevant outcomes that can also be important for evaluating upper limb motor recovery, such as range of motion (ROM), muscle strength and quality of life (QoL), is still weak due to the low quality of evidence.<sup>20</sup> Additionally, previous studies have primarily used non-immersive VR. The use, effectiveness and safety of immersive VR in poststroke motor rehabilitation are less studied and thus are less understood.<sup>20–26</sup> Moreover, users’ perceptions of a technology, such as its ease of use and usefulness, can influence their attitudes toward the technology and ultimately impact their acceptance of the technology.<sup>8,27–39</sup> In the context of the use of technology in poststroke motor rehabilitation, stroke patients’ perceptions of VR-based rehabilitation programmes can also influence their acceptance of and adherence to the programmes; however, research on this topic has been limited.

In view of the aforementioned research gaps, we conducted a proof-of-concept and feasibility randomized controlled trial (RCT) to examine the effectiveness, safety and patients’ perceptions of an immersive VR-based exercise system developed to support the performance of upper limb exercises in poststroke patients.

## Methods

### Study design and participants

A proof-of-concept, parallel-group, double-blind RCT was conducted at a public hospital in China (Dingzhou People’s

Hospital, Hebei) from September 2021 to June 2022. Participants were poststroke patients recruited from the hospital’s rehabilitation medicine department. Eligible participants were aged 19–75 years; had a first-ever unilateral stroke, as confirmed from their computed tomography or magnetic resonance imaging records; were experiencing motor impairments in one of the upper limbs, with the affected arm in Brunnstrom stage 3, 4, or 5 of stroke recovery; had at least 10° of active range of motion (AROM) in the shoulder and elbow of the affected arm; were able to maintain autonomous upright seating for at least 45 minutes; had a normal or corrected-to-normal vision; and had normal hearing. Patients were excluded if they had injuries or other health conditions that restricted their upper limb mobility, unilateral spatial neglect, unstable medical conditions, a history of seizures or epilepsy, communication difficulty, mood instability or were involved in any other ongoing investigational drug studies.

The trial was approved by the Human Research Ethics Committee of The University of Hong Kong-Shenzhen Hospital (approval number: [2021]137) and the Dingzhou People’s Hospital (approval number: hx-2021-04) in China and has been registered in the Chinese Clinical Trial Registry (registration number: ChiCTR2100047150). All of the participants provided written informed consent before participating in the study.

### Procedures

Doctors and physical therapists in the rehabilitation medicine department first introduced the study to their stroke inpatients. Patients who were interested in the study were invited to a screening session, during which they tried the study’s VR devices and their eligibility was determined. The researcher (JC) explained the study’s details to the patients initially determined to be eligible, confirmed their eligibility and obtained their written informed consent. Subsequently, the researcher collected data on the participants’ demographic and clinical characteristics, and two outcome assessors (YZ and TH) conducted baseline assessments.

The participants were then randomized to one of the two study groups: the intervention group, in which the immersive VR-based exercise system was used to perform upper limb exercises (35 minutes/day, 6 days/week for 2 weeks), or the control group, in which commercial games were played (35 minutes/day, 6 days/week for 2 weeks). The dose and frequency of the VR exercise or game playing were suggested by physical therapists in accordance with the dose and frequency of the physical therapy at the hospital (30 minutes/day, 6 days/week). The study was conducted in a separate therapy room (approximately 25 m<sup>2</sup>) in the department. The outcome measures were assessed at baseline and at 1-week and 2-week follow-ups.

All of the participants continued their stroke rehabilitation therapy provided by the hospital during the study

period, including physical therapy (30 minutes/day), functional electrical stimulation (20 to 80 minutes/day), occupational therapy (20 minutes/day), aerobic exercises (15 to 85 minutes/day), acupuncture (30 minutes/day), soft robotic glove-assisted hand training (15 minutes/day), intermittent pneumatic compression therapy for lower extremity (30 minutes/day), facial massage (15 minutes/day), speech therapy (20 minutes/day), transcranial magnetic stimulation (20 minutes/day) and equipment-assisted standing training (30 minutes/day). The above therapy was delivered 6 days per week.

### Randomization and masking

The participants were randomized 1:1 to either the intervention or the control group, using permuted blocks of sizes four and six. Randomization was performed using a computer program. A researcher (TC) who was not involved in the study performed the randomization after the baseline assessments had been completed. YZ and TH (the two outcome assessors) and the participants were blinded to group allocation. As JC (the onsite researcher) had to administer the intervention, researcher blinding was not possible.

### Intervention and control conditions

**Intervention condition.** The participants randomized to the intervention group were instructed to perform upper limb exercises using the immersive VR-based exercise system. The design of the system adhered to the established learning principles of poststroke rehabilitation.<sup>3,5,40,41</sup> First, the exercise system focused on training the fundamental upper limb movements essential to daily activities (e.g., shoulder flexion). The tasks involved in the exercise system were those that patients were familiar with in daily life, such as lifting weights (in our case, a dumbbell). Second, the exercise system provided repetitive exercises of the upper limb to facilitate the development of normal movement pattern. Third, the exercise system provided visual (e.g., balloon popping into pieces), audio (e.g., sheep bleating) and haptic (e.g., fish rod vibrating) feedback to inform patients' success in performing the tasks and assist their motor learning. Fourth, the exercise system incorporated a progressive difficulty level for tasks, to suit the heterogeneous conditions and abilities of patients. The difficulty levels could be quantified as the number of repetitions of movements, degrees and distances, allowing patients setting quantifiable goals of performing tasks. Fifth, the tasks were designed as engaging cartoon games to increase patients' enthusiasm to participate actively in the exercises.

The exercise system included five games, one each for shoulder flexion and abduction (*Dumbbell lifting*), elbow flexion (*Fishing*), forearm pronation and supination

(*Sheep whacking*), wrist flexion and extension (*Apple picking*) and reaching exercises (*Balloon popping*), which were based on the Graded Repetitive Arm Supplementary Program (<https://neurorehab.med.ubc.ca/grasp/>)<sup>42-44</sup> and PhysioTherapy eXercises (<https://www.physiotherapyexercises.com/>). The games were recently preliminarily applied and their feasibility demonstrated in healthy older adults.<sup>30</sup> The details of the five exercise games are presented below.

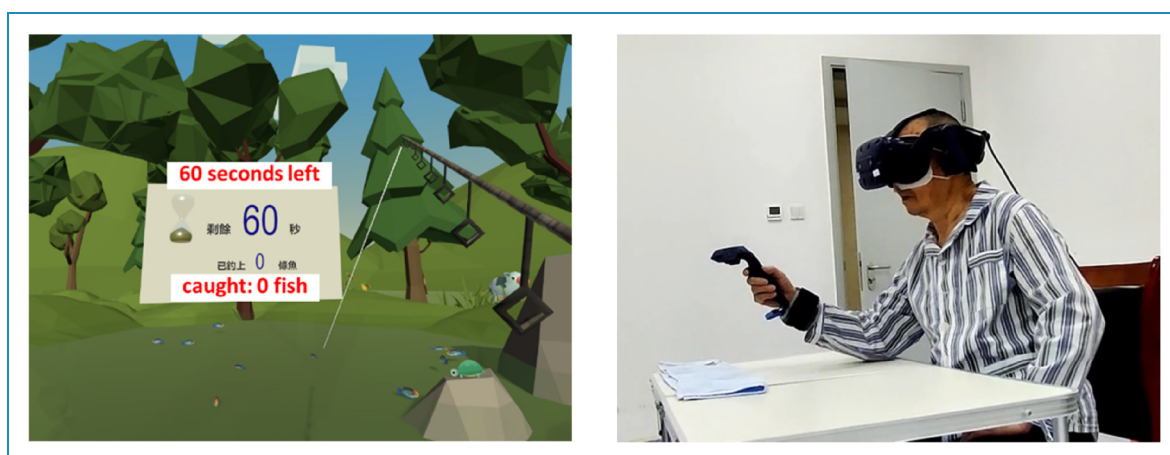
**Dumbbell lifting.** This game included two modes: shoulder flexion exercise and shoulder abduction exercise (Figure 1). The participants held the controller (presented as a dumbbell in VR) and needed to (1) start with their arm hanging naturally to the side of their body, elbow fully extended and palm facing the side of the body; then (2) lift the controller up and out to the front (for flexion) or to the side (for abduction) until it reached shoulder level, keeping the elbow straight; and then (3) hold the controller at the shoulder level for 1 to 3 seconds, depending on the difficulty level. A longer required time for holding the controller indicated a higher difficulty level than a shorter required time for holding the controller. One trial lasted 1 minute, during which each participant was required to repeatedly lift the controller by following the above-described steps as many times as possible. For each mode, four trials were performed by each participant.

**Fishing.** This game was designed as an elbow flexion exercise (Figure 2). The participants were in front of a pond in VR in which fish were swimming. The participants needed to (1) hold the controller (presented as a fishing rod in VR) and wait until it started vibrating, indicating that a fish had been caught; then (2) flex their elbow and raise the controller toward the shoulder, keeping the elbow still on the real table; and then (3) hold the position for 1 to 3 seconds, depending on the difficulty level, to pull the fish out of the water. A longer required time for holding the controller indicated a higher difficulty level than a shorter required time holding the controller. One trial lasted 1 minute, during which each participant was required to catch as many fish as possible. Each participant performed four trials.

**Sheep whacking.** This game was designed as a forearm pronation and supination exercise (Figure 3). The participants were in front of two holes in VR, holding the controller (presented as a hammer in VR). A sheep alternately popped up from one of the two holes. The participants needed to whack the sheep back into the hole by pronating or supinating their forearm, with the elbow kept still on the real table. There were four difficulty levels: the participants needed to pronate and supinate the forearm to 15°, 30°, 45° and 60°; more required degrees of movement indicated higher difficulty levels than fewer required degrees of movement. One trial lasted 1 minute, during which each



**Figure 1.** Dumbbell lifting in virtual reality (left) for shoulder exercises (right).



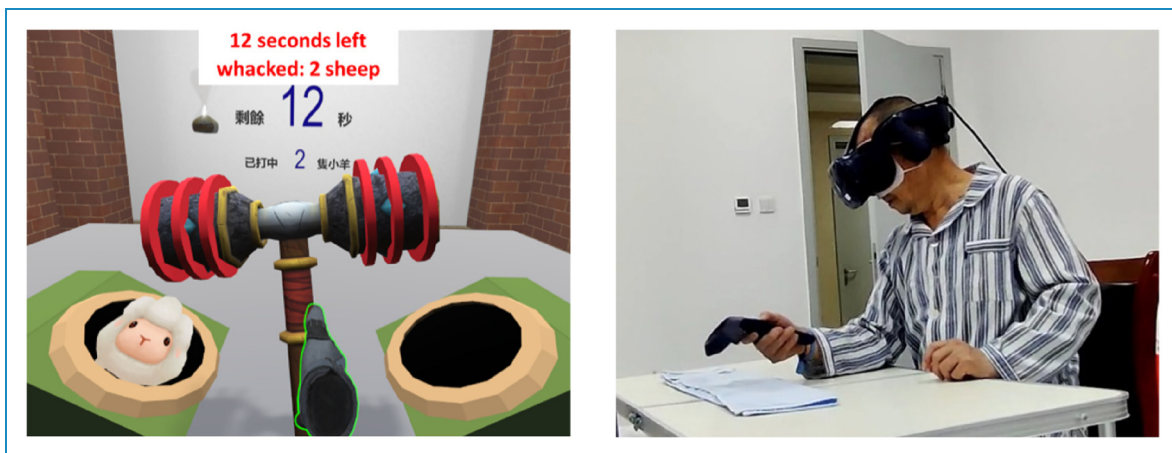
**Figure 2.** Fishing in virtual reality (left) for elbow exercise (right).

participant was required to whack as many sheep as possible. Each participant performed four trials.

**Apple picking.** This game was designed as a wrist flexion and extension exercise (Figure 4). The participants were in front of an apple tree and a stump in VR. The participants needed to hold the controller (presented as a bird in VR) and extend their wrist to control the upward flight of the bird to pick a red apple from the tree, then flex their wrist to control the downward flight of the bird and drop the apple on the stump. Immediately after this, another apple appeared on the tree, and the participants needed to repeat the process. There were four difficulty levels: the participants needed to flex and extend their wrist to 15°, 30°, 45° and 60°; more required degrees of movement indicated higher difficulty levels than less required degrees of movement. One trial lasted 1 minute, during which each participant was required to pick as many apples as possible. Each participant performed four trials.

**Balloon popping.** This game was designed as a reaching exercise for the overall upper limb (Figure 5), which combined the movements of shoulder flexion and extension, and elbow flexion and extension. The participants were in front of a virtual table, on which a balloon appeared on the affected side of the participant. The participants needed to hold the controller (presented as a hand in VR) and reach their hand toward the balloon and pop it by pressing a trigger button on the controller using their index finger. Immediately after this, another balloon appeared, and the participants needed to repeat the process. The game had four modes: reaching toward the front, reaching toward the side at 45° and 90° and reaching upward (toward a balloon). There were four difficulty levels within each mode, depending upon the distance between the balloons and the participants. Balloons randomly appeared within a distance of 0%–25%, 0%–50%, 0%–75% and 0%–100% of the participants' maximum reaching distance. A longer distance indicated a higher difficulty level than a shorter distance. Each participant's maximum reaching distance was measured before the game. One trial lasted 1 minute,





**Figure 3.** Sheep whacking in virtual reality (left) for forearm exercises (right).

during which each participant was required to pop as many balloons as possible. Each participant performed four trials.

Participants in the intervention group started with *Dumbbell lifting*, followed by *Fishing*, *Sheep whacking*, *Apple picking* and *Balloon popping*. The difficulty level of the games started from the easiest level. After each trial, the researcher asked the participants how difficult they felt the exercise was; if the participants reported the trial was easy, the difficulty level was increased by one; if the participants reported the trial was acceptable or difficult to perform, the difficulty level remained unchanged. After consultations with physical therapists, the participants who had been cleared to perform strengthening exercises used weights (0.25, 0.5 or 1 kg, based on the level of their motor function) placed on the wrist for the shoulder, elbow, forearm, and reaching exercises or on the hand for the wrist exercises. Rest during exercises was allowed to avoid fatigue. The participants were instructed to stretch their upper limbs during the rest to avoid increased muscle spasticity. Guided by the physical therapists, the researcher held the participants' joints to ensure their sitting balance and avoid abnormal movements (e.g., shrugging the shoulders) when needed.

A HMD (HTC Vive Pro) was used for navigation in virtual environments, and a wireless hand-held controller was used for interaction (e.g. lifting the dumbbell). Unity 2018.4.24f1 (Unity Technologies, USA) and VotanicXR (version 2020, Votanic Ltd., HK, China) were used to develop the exercise system, which was run on a Dell G7 laptop computer. An elastic bandage was used to maintain the hand-held controller on the hand if the participants were not able to independently hold the hand-held controller with their affected hand. A table was used to support the elbow during the elbow and forearm exercises. A sponge foam box placed on the table was used to support the wrist exercises.

**Control condition.** Immersive VR-based commercial games obtained from Steam (<https://store.steampowered.com/>) were

used as a sham VR programme, in order to blind the participants and outcome assessors to the group allocation. The games used in the study required upper limb movements (it is impossible to avoid upper limb movements with immersive VR devices) but do not require specific movement patterns, and thus, they were not directly aligned with the exercises provided by the system we developed. The participants randomized to the control group play the games for entertainment, rather than developing normal movement patterns. The following games were used after their safety and playability had been assessed by the doctors, physical therapists and researcher of this study: *Chocolat Rush*, *Zooma*, *Fruit Ninja VR*, *Koorring Wonderland*, and *Space Slurpies* (details of the games are presented in Appendix 1). These games are single-player, casual and supported on HTC Vive Pro, and do not involve highly active movements, complex operations or elements that may induce discomfort. After the researcher explained the rules of the games, the participants were told to select several games they were interested in and play them at their style and pace, during which each game was repeated approximately two to four times. The participants rested and stretched their upper limbs when needed to avoid fatigue and increased muscle spasticity. The researcher held the participant's joints to ensure their sitting balance and avoid abnormal movements when needed, with the guidance of the physical therapists.

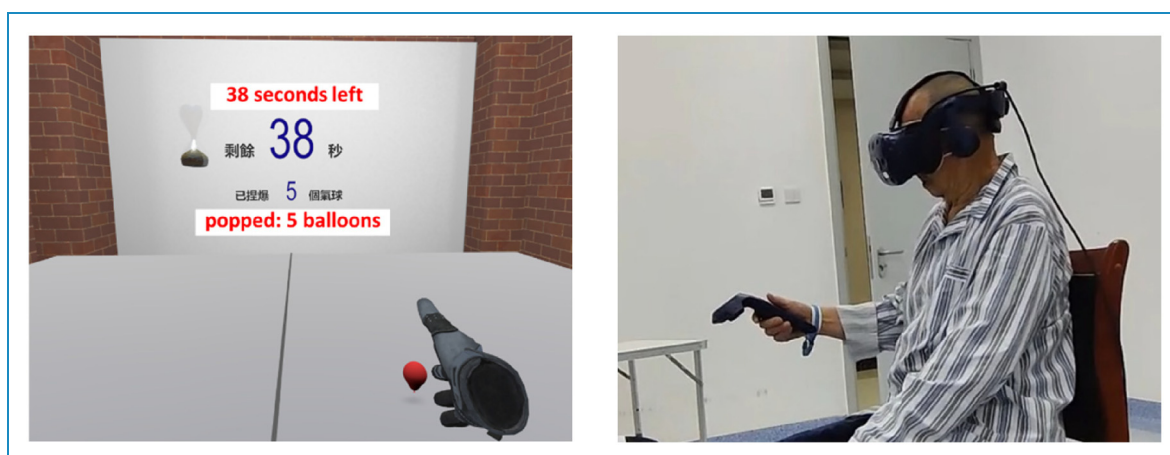
The control participants used the same HMD, hand-held controller, laptop computer and elastic bandage as the intervention participants.

## Outcomes

**Baseline demographic and clinical characteristics.** The baseline demographic and clinical characteristics measured were age, sex, education level, stroke type, time since stroke onset, stroke-affected side, dominant hand before stroke, Brunnstrom stage of the affected arm, comorbidities,



**Figure 4.** Apple picking in virtual reality (left) for wrist exercises (right).



**Figure 5.** Balloon popping in virtual reality (left) for reaching exercise (right).

ongoing stroke rehabilitation therapy prescribed by doctors and previous experience of using a computer and VR.

*Outcomes used to evaluate the effectiveness of the immersive VR-based exercise system.* The effectiveness outcomes were as follows:

- Upper limb motor function measured using the FMA-UE scale, which contains four subscales: shoulder/elbow/forearm, wrist, hand and coordination/speed scales. The total score ranges from 0 to 66, with higher scores indicating better motor function than lower scores.<sup>45</sup>
- AROM of shoulder flexion and abduction, elbow flexion, forearm pronation and supination, and wrist flexion and extension, measured using a goniometer.<sup>46,47</sup>
- Strength of shoulder flexors and abductors, elbow flexors, and wrist flexors and extensors measured using a hand-held muscle dynamometer (model 01165; Lafayette Instrument Company, USA).<sup>48–51</sup>
- Arm and hand motor ability measured using the Wolf Motor Function Test (WMFT), which consists of 15 timed tasks, with up to 120 seconds allowed per task.<sup>52</sup> The task completion times of the 15 tasks were averaged, with a shorter completion time indicating better motor ability than a longer completion time. The quality of movement while performing the tasks was rated on a 6-point scale, ranging from 0 (does not attempt to perform the task) to 5 (movement appears to be normal). The quality of movement scores of the 15 tasks were averaged, with higher scores indicating better motor ability than lower scores.
- Perceived upper limb motor function of the affected side measured using an 11-point rating scale ranging from 0 (very poor) to 10 (excellent).
- QoL measured using the EQ-5D-5L questionnaire (<https://euroqol.org/>), which measures five dimensions of health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension is rated on a 5-point

**Table 1.** Perception outcomes and measurement items.

Outcomes	Measurement items
Perceived usefulness (PU)	PU1: Using the immersive VR-based exercise system to perform upper limb exercises would improve your upper limb motor function.
	PU2: Using the immersive VR-based exercise system to perform upper limb exercises would increase exercise efficiency.
	PU3: Using the immersive VR-based exercise system to perform upper limb exercises would enhance exercise effectiveness.
	PU4: You found the immersive VR-based exercise system to be useful to perform upper limb exercises.
Perceived ease of use (PEOU)	PEOU1: Learning to use the immersive VR-based exercise system was easy for you.
	PEOU2: You found it easy to get the immersive VR-based exercise system to do what you wanted to do.
	PEOU3: It was easy to become skillful at using the immersive VR-based exercise system.
	PEOU4: You found the immersive VR-based exercise system easy to use.
Attitude (ATT)	ATT1: Using the immersive VR-based exercise system to perform upper limb exercises is a good idea.
	ATT2: Using the immersive VR-based exercise system to perform upper limb exercises is a wise idea.
	ATT3: You like the idea of using the immersive VR-based exercise system to perform upper limb exercises.
	ATT4: Using the immersive VR-based exercise system to perform upper limb exercises was a pleasant experience.
Intrinsic motivation (IM)	IM1: You found using the immersive VR-based exercise system to perform upper limb exercises to be enjoyable.
	IM2: The actual process of using the immersive VR-based exercise system to perform upper limb exercises was pleasant.
	IM3: You had fun using the immersive VR-based exercise system to perform upper limb exercises.
Satisfaction (SAT)	SAT1: You are satisfied with the use of the immersive VR-based exercise system for performing upper limb exercises.
	SAT2: You are satisfied with the immersive VR-based exercise system.
Behavioural intention (BI)	BI1: You intend to continue using the immersive VR-based exercise system to perform upper limb exercises.
	BI2: You plan to continue using the immersive VR-based exercise system to perform upper limb exercises.

VR, virtual reality.

scale ranging from 1 (no problem) to 5 (unable to perform/ extreme problem). The responses were then summarized as five-digit codes (e.g. 11235), and the codes were transformed into a single number (i.e. an index value) using the Chinese population tariff to present a specific health status.<sup>53</sup>

*Outcomes used to evaluate the safety of the immersive VR-based exercise system.* The following three outcomes were used to examine the safety of the exercise system:

(i) discomfort, (ii) pain in the upper limb, and (iii) muscle spasticity in the upper limb, all due to the use of the exercise system. Discomfort was recorded based on the participants' self-report and the researcher's observation during each session of VR exercise. Pain was measured using an 11-point visual analogue scale ranging from 0 (no pain) to 10 (worst possible pain), with higher scores indicating more severe pain than lower scores. Muscle spasticity was measured using the modified Ashworth Scale,

ranging from 0 to 4 with 6 possible levels (i.e. 0, 1, 1+, 2, 3, 4), with higher scores indicating more severe muscle spasticity than lower scores.<sup>49</sup>

*Outcomes used to evaluate patients' perceptions of the immersive VR-based exercise system.* In the last follow-up assessment, a questionnaire was used to measure the intervention participants' perceived usefulness of the exercise system, perceived ease of use of the exercise system, attitudes toward the exercise system, intrinsic motivation for using the exercise system, satisfaction with the exercise system and intention to use the exercise system. Table 1 presents the perception outcomes and their measurement items, which were rated on a 7-point Likert scale ranging from 1 (very strongly disagree) to 7 (very strongly agree).<sup>36,37,54,55</sup>

### Statistical analysis

Between-group differences in participant characteristics and outcome variables at baseline were assessed using independent *t*-tests and chi-square/Fisher exact tests. Mean changes in the effectiveness outcomes over time within the groups and between the two groups were analysed using linear mixed models with a scaled identity covariance structure. The linear mixed model technique can accommodate all available data of participants if missing data appear, without dropping the entire data of the participants from the analysis.<sup>56</sup> The outcomes were modelled in terms of the time effect (1 week vs. baseline and 2 weeks vs. baseline) and treatment condition effect (intervention vs. control). Intention-to-treat analysis was first performed by including all participants in the study group to which they had been assigned. We assumed that the missing data were missing at random and no imputation was adopted for the missing data.<sup>57</sup> We also conducted a per-protocol analysis which included only participants who completed the 2-week trial. The safety and patient perception outcomes were reported descriptively. IBM SPSS Statistics (version 25) was used for all of the statistical analyses. A two-sided  $p < 0.05$  was considered statistically significant.

## Results

### Study characteristics

From September 2021 to June 2022, 84 patients were screened. Fifty of them were eligible and willing to participate, and were thus enrolled in the study and randomized into the intervention and control groups (Figure 6). Two intervention participants were lost to follow-up at week 1: one was discharged, and another was unable to arrange a time to participate in the study; nine intervention participants were lost to follow-up at week 2: seven were

discharged, and two did not feel well during hospitalization. Seven control participants were lost to follow-up at week 1: three were discharged; three lost interest and one participant's family members thought the participant was too tired to attend the study; four control participants were lost to follow-up at week 2, because they were discharged. Baseline characteristics and outcome measures were balanced between the groups, except that the intervention group had a significant higher prevalence of diabetes than the control group ( $p < 0.001$ ) (Tables 2 and 3).

### Effectiveness outcomes

*Intention-to-treat analysis.* The results of the intention-to-treat analyses of effectiveness outcomes are shown in Table 4.

Regarding upper limb motor function measured using the FMA-UE, a significant increase was observed in the coordination/speed score after 1 week in the intervention group (0.51, 95% confidence interval [CI]: 0.05 to 0.97,  $p = 0.03$ ), but no significant difference was noted compared with the control group ( $p = 0.94$ ). No other within-group and between-group differences were observed.

For AROM, significant increases were observed in the AROM of shoulder flexion after 2 weeks in both the intervention (19.71, 95% CI: 3.33 to 36.09,  $p = 0.02$ ) and control groups (17.17, 95% CI: 0.79 to 33.55,  $p = 0.04$ ), but no significant between-group difference was observed ( $p = 0.83$ ). Significant increases were also observed in the AROM of shoulder abduction after 1 week (16.69, 95% CI: 3.36 to 30.03,  $p = 0.02$ ) and 2 weeks (29.21, 95% CI: 13.81 to 44.62,  $p < 0.001$ ) in the intervention group and after 2 weeks (16.81, 95% CI: 1.41 to 32.22,  $p = 0.03$ ) in the control group, but no significant between-group differences were found at 1 week ( $p = 0.55$ ) or 2 weeks ( $p = 0.26$ ). No other within-group and between-group differences were observed.

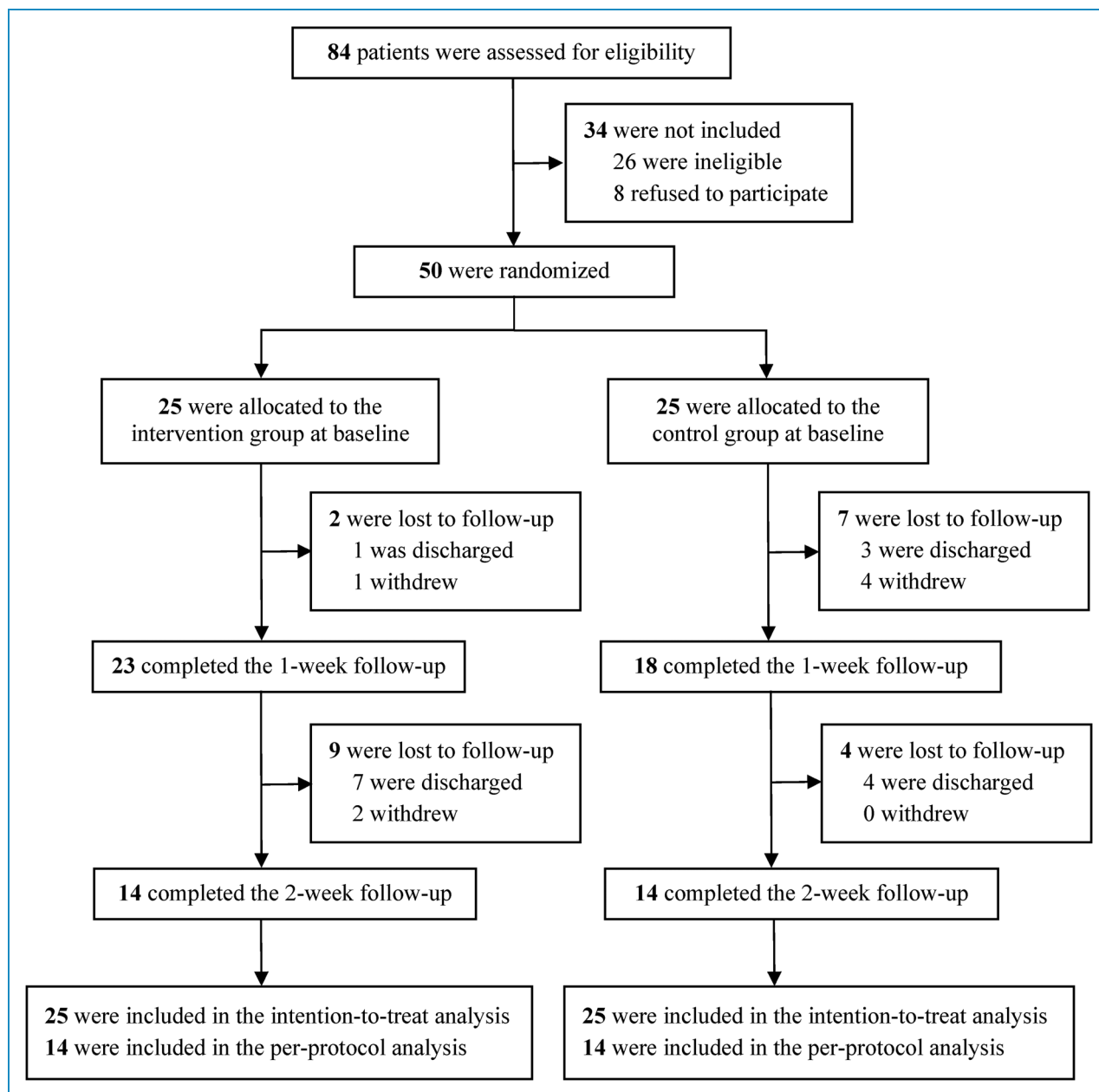
There were no within-group or between-group differences in strength or arm and hand motor ability, as measured using the WMFT.

Regarding perceived upper limb motor function, significant increases were observed after 2 weeks (1.78, 95% CI: 0.67 to 2.89,  $p = 0.002$ ) in the intervention group and after 1 week (1.33, 95% CI: 0.31 to 2.36,  $p = 0.01$ ) and 2 weeks (1.64, 95% CI: 0.53 to 2.75,  $p = 0.004$ ) in the control group, but no significant between-group differences were found at 1 week ( $p = 0.39$ ) or 2 weeks ( $p = 0.86$ ).

For QoL, significant increases were observed after 1 week (0.14, 95% CI: 0.02 to 0.25,  $p = 0.02$ ) and 2 weeks (0.18, 95% CI: 0.05 to 0.30,  $p = 0.01$ ) in the control group, but no significant between-group differences were found at 1 week ( $p = 0.55$ ) or 2 weeks ( $p = 0.58$ ).

*Per-protocol analysis.* The results of the per-protocol analyses of effectiveness outcomes are shown in Table 5.





**Figure 6.** Study flow diagram.

Regarding upper limb motor function measured using the FMA-UE, a significant increase was observed in the coordination/speed score after 2 weeks in the intervention group (0.64, 95% CI: 0.06 to 1.23,  $p=0.03$ ), but there was no significant between-group difference ( $p=0.73$ ). No other within-group and between-group differences were observed.

For AROM, significant increases were observed in the AROM of shoulder flexion after 1 week (17.03, 95% CI: 1.57 to 32.48,  $p=0.03$ ) and 2 weeks (28.36, 95% CI: 12.90 to 43.82,  $p<0.001$ ) in the intervention group, but no significant between-group differences were observed at 1 week ( $p$

$=0.33$ ) or 2 weeks ( $p=0.15$ ). Significant increases were also observed in the AROM of shoulder abduction after 1 week (17.57, 95% CI: 1.86 to 33.29,  $p=0.03$ ) and 2 weeks (36.24, 95% CI: 20.53 to 51.95,  $p<0.001$ ) in the intervention group, and the improvement in AROM in the intervention group was significantly higher than that in the control group at 2 weeks ( $p=0.04$ ).

No within-group and between-group differences were observed in strength and arm and hand motor ability, as measured using the WMFT.

Regarding perceived upper limb motor function, significant increases were observed after 2 weeks in the

**Table 2.** Baseline characteristics of the participants.

Characteristics	Intervention (n = 25)	Control (n = 25)	p value for between-group difference
Age (years)			
Mean (SD)	56.4 (12.9)	59.1 (10.2)	0.41
Median (range)	58.0 (28.0–75.0)	60.0 (34.0–75.0)	
Sex, n (%)			
Male	16 (64.0%)	17 (68.0%)	0.77
Female	9 (36.0%)	8 (32.0%)	
Education received (years), n (%)			
≤6	0 (0.0%)	1 (4.0%)	0.27
6–9	24 (96.0%)	19 (76.0%)	
>9	1 (4.0%)	5 (20.0%)	
Stroke type, n (%)			
Ischaemic	17 (68.0%)	16 (64.0%)	0.77
Haemorrhagic	8 (32.0%)	9 (36.0%)	
Time since stroke onset at baseline (days)			
Mean (SD)	36.1 (24.2)	46.0 (74.2)	0.53
Median (range)	34.0 (7.0–107.0)	19.0 (5.0–357.0)	
Stroke-affected side, n (%)			
Left	13 (52.0%)	13 (52.0%)	>0.99
Right	12 (48.0%)	12 (48.0%)	
Dominant hand, n (%)			
Left	1 (4.0%)	4 (16.0%)	0.35
Right	24 (96.0%)	21 (84.0%)	
Brunnstrom stage of stroke recovery for the affected arm, n (%)			
Stage 3	14 (56.0%)	11 (44.0%)	0.69
Stage 4	6 (24.0%)	8 (32.0%)	
Stage 5	5 (20.0%)	6 (24.0%)	

(continued)

Table 2. Continued.

Characteristics	Intervention (n = 25)	Control (n = 25)	p value for between-group difference
Comorbidities, n (%)			
Hypertension	20 (80.0%)	21 (84.0%)	>0.99
Diabetes	13 (52.0%)	1 (4.0%)	<0.001
Hyperlipidaemia	5 (20.0%)	5 (20.0%)	>0.99
Vascular disease	4 (16.0%)	5 (20.0%)	>0.99
Heart disease	4 (16.0%)	4 (16.0%)	>0.99
Musculoskeletal disorder	1 (4.0%)	1 (4.0%)	>0.99
No comorbidities	0 (0.0%)	1 (4.0%)	>0.99
Conventional therapeutic modalities, n (%)			
Physical therapy	25 (100.0%)	23 (92.0%)	0.49
Functional electric stimulation	25 (100.0%)	23 (92.0%)	0.49
Occupational therapy	25 (100.0%)	21 (84.0%)	0.11
Aerobic exercises	23 (92.0%)	21 (84.0%)	0.67
Acupuncture	23 (92.0%)	21 (84.0%)	0.67
Soft robotic glove-assisted hand training	16 (64.0%)	10 (40.0%)	0.09
Intermittent pneumatic compression therapy for lower extremity	6 (24.0%)	5 (20.0%)	0.73
Facial massage	3 (12.0%)	5 (20.0%)	0.70
Speech therapy	2 (8.0%)	2 (8.0%)	>0.99
Transcranial magnetic stimulation	1 (4.0%)	3 (12.0%)	0.61
Equipment-assisted standing training	2 (8.0%)	1 (4.0%)	>0.99
Experience of using a computer, n (%)			0.19
Never	17 (68.0%)	19 (76.0%)	
Rarely	2 (8.0%)	5 (20.0%)	
Sometimes	3 (12.0%)	0 (0.0%)	
Often	3 (12.0%)	1 (4.0%)	
Always	0 (0.0%)	0 (0.0%)	

(continued)

Table 2. Continued.

Characteristics	Intervention (n = 25)	Control (n = 25)	p value for between-group difference
Experience of using VR, n (%)			>0.99
Never	24 (96.0%)	24 (96.0%)	
Rarely	1 (4.0%)	1 (4.0%)	
Sometimes	0 (0.0%)	0 (0.0%)	
Often	0 (0.0%)	0 (0.0%)	
Always	0 (0.0%)	0 (0.0%)	
Total duration of VR exercises/game playing (minutes) <sup>a</sup>			
Mean (SD)	296.04 (132.97)	267.40 (162.56)	0.50
Median	350.0	350.0	
Range	35.0–420.0	35.0–420.0	

<sup>a</sup>One intervention participant did not receive virtual reality exercises because the participant withdrew from the study immediately after enrolment. SD, standard deviation; VR, virtual reality.

intervention group (2.14, 95% CI: 0.90 to 3.39,  $p=0.001$ ) and control group (1.29, 95% CI: 0.04 to 2.53,  $p=0.04$ ), but no significant between-group difference was found at 2 weeks ( $p=0.34$ ).

For QoL, a significant increase was observed after 2 weeks (0.15, 95% CI: 0.02 to 0.28,  $p=0.02$ ) in the control group, but there was no significant between-group difference at 2 weeks ( $p=0.57$ ). Appendix 2 summarizes the distribution of the participants' responses to the EQ-5D-5L.

### Safety outcomes

One intervention participant experienced dizziness from using VR. Two intervention participants and one control participant reported eye strain due to wearing the HMD. One intervention participant experienced fatigue during VR exercise. Two intervention participants reported pain in the upper limb during VR exercise.

Table 6 presents the distribution of the participants' ratings for pain in their upper limb due to the use of the exercise system at 1 and 2 weeks. More than half of the participants in both groups perceived no pain in the upper limb, and several participants reported mild to moderate levels of pain, with ratings ranging from 2 to 6.

Table 7 shows the distribution of the muscle spasticity of the biceps, triceps, forearm pronators, forearm

supinators, wrist flexors and wrist extensors for both groups at 1 and 2 weeks. Almost all the participants had no increase in spasticity (level 0) or mild spasticity (level 1 or 1+), and no severe spasticity was reported.

### Patient perception outcomes

Of the 25 intervention participants, 21 completed the patient perception questionnaire. The mean values of the ratings (ranging from 1 to 7, with higher values indicating more positive perceptions than lower values) for perceived usefulness, perceived ease of use, attitude, intrinsic motivation, satisfaction and behavioural intention were 6.20 (standard deviation [SD]=0.69), 6.33 (SD=0.80), 6.25 (SD=0.74), 6.37 (SD=0.75), 6.36 (SD=0.69) and 6.31 (SD=0.84), respectively. These results indicate that the participants' overall perceptions of the exercise system were highly positive.

## Discussion

### Principal findings

This study developed an immersive VR-based exercise system and evaluated its use for upper limb motor rehabilitation in poststroke patients from multiple perspectives. Regarding its effectiveness, the intention-to-treat analyses showed that there were no significant differences in upper limb motor function (FMA-UE), upper limb AROM,



**Table 3.** Outcomes at baseline.

Outcome	Intervention (n = 25)	Control (n = 25)	p value for between-group difference
Upper limb motor function (Fugl-Meyer Assessment for Upper Extremity)			
Shoulder/elbow/forearm	20.04 (9.44)	20.36 (8.63)	0.90
Wrist	3.40 (3.72)	3.08 (3.20)	0.75
Hand	7.56 (4.74)	7.12 (4.76)	0.75
Coordination/speed	4.36 (0.91)	4.24 (0.78)	0.62
Total	35.36 (17.16)	34.80 (15.20)	0.90
Active range of motion (°)			
Shoulder flexion	98.60 (26.47)	96.64 (26.49)	0.80
Shoulder abduction	81.81 (20.95)	84.85 (24.97)	0.64
Elbow flexion	107.08 (25.68)	104.84 (20.34)	0.73
Forearm pronation	48.79 (35.37)	51.83 (35.82)	0.76
Forearm supination	47.20 (33.98)	50.88 (31.83)	0.69
Wrist flexion	20.75 (21.64)	23.60 (20.69)	0.64
Wrist extension	21.81 (19.21)	26.56 (23.81)	0.44
Strength (N)			
Shoulder flexors	50.24 (29.23)	42.55 (19.05)	0.28
Shoulder abductors	53.54 (29.19)	43.44 (18.79)	0.15
Elbow flexors	47.54 (29.18)	42.80 (23.37)	0.53
Wrist flexors	13.38 (18.93)	12.37 (14.68)	0.83
Wrist extensors	21.67 (25.59)	15.34 (17.12)	0.31
Arm and hand motor ability (Wolf Motor Function Test)			
Task completion time (seconds)	50.40 (35.88)	52.37 (32.83)	0.84
Quality of movement (score)	2.59 (1.32)	2.35 (1.02)	0.48
Perceived upper limb motor function	3.72 (1.51)	4.00 (1.55)	0.52
Quality of life (EQ-5D-5L)	0.57 (0.21)	0.55 (0.24)	0.66

**Table 4.** Effects of the virtual reality-based exercise intervention on effectiveness outcomes based on the intention-to-treat analyses.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	n	Mean (95% CI)	Change from baseline	
Upper limb motor function (Fugl-Meyer Assessment for Upper Extremity)								
Shoulder/elbow/forearm								
Baseline	25	20.04 (3.69 to 36.40)	NA	NA	25	20.36 (4.01 to 36.72)	NA	NA
Week 1	23	22.87 (2.66 to 43.08)	2.83 (-2.33 to 7.99)	0.28	18	23.33 (4.67 to 41.99)	2.97 (-2.55 to 8.50)	0.29
Week 2	14	21.86 (1.99 to 41.72)	1.82 (-4.15 to 7.78)	0.55	14	23.57 (3.70 to 43.44)	3.21 (-2.75 to 9.18)	0.29
Wrist								
Baseline	25	3.40 (1.94 to 4.86)	NA	NA	25	3.08 (1.62 to 4.54)	NA	NA
Week 1	23	3.78 (2.26 to 5.31)	0.38 (-1.73 to 2.49)	0.72	18	4.17 (2.45 to 5.89)	1.09 (-1.17 to 3.34)	0.34
Week 2	14	3.43 (1.48 to 5.38)	0.03 (-2.41 to 2.47)	0.98	14	4.07 (2.12 to 6.02)	0.99 (-1.45 to 3.43)	0.42
Hand								
Baseline	25	7.56 (5.65 to 9.48)	NA	NA	25	7.12 (5.21 to 9.04)	NA	NA
Week 1	23	8.39 (6.40 to 10.39)	0.83 (-1.94 to 3.60)	0.55	18	8.00 (5.74 to 10.26)	0.88 (-2.08 to 3.84)	0.56
Week 2	14	7.43 (4.87 to 9.99)	-0.13 (-3.33 to 3.06)	0.94	14	8.36 (5.80 to 10.92)	1.24 (-1.96 to 4.43)	0.45
Coordination/speed								
Baseline	25	4.36 (4.04 to 4.68)	NA	NA	25	4.24 (3.92 to 4.56)	NA	NA
Week 1	23	4.87 (4.54 to 5.20)	0.51 (0.05 to 0.97)	0.03	18	4.72 (4.35 to 5.10)	0.48 (-0.01 to 0.98)	0.06
Week 2	14	4.86 (4.43 to 5.28)	0.50 (-0.04 to 1.03)	0.07	14	4.71 (4.29 to 5.14)	0.47 (-0.06 to 1.01)	0.08

(continued)

Table 4. Continued.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	Mean (95% CI)	Change from baseline	p value	
Total								
Baseline	25	35.36 (20.57 to 50.15)	NA	NA	25	34.80 (20.01 to 49.59)	NA	NA
Week 1	23	39.91 (24.99 to 54.84)	4.55 (-5.08 to 14.19)	0.35	18	40.22 (24.86 to 55.58)	5.42 (-4.88 to 15.73)	0.30
Week 2	14	37.57 (21.65 to 53.49)	2.21 (-8.92 to 13.34)	0.70	14	40.71 (24.80 to 56.63)	5.91 (-5.22 to 17.04)	0.30
Active range of motion (°)								
Shoulder flexion								
Baseline	25	98.60 (70.66 to 126.54)	NA	NA	25	96.64 (68.70 to 124.58)	NA	NA
Week 1	23	109.64 (81.55 to 137.72)	11.04 (-3.14 to 25.21)	0.13	18	104.45 (75.85 to 133.04)	7.81 (-7.36 to 22.97)	0.31
Week 2	14	118.31 (89.07 to 147.55)	19.71 (3.33 to 36.09)	0.02	14	113.81 (84.57 to 143.05)	17.17 (0.79 to 33.55)	0.04
Shoulder abduction								
Baseline	25	81.81 (72.58 to 91.04)	NA	NA	25	84.85 (75.62 to 94.08)	NA	NA
Week 1	23	98.51 (88.89 to 108.13)	16.69 (3.36 to 30.03)	0.02	18	95.63 (84.75 to 106.51)	10.78 (-3.49 to 25.04)	0.14
Week 2	14	111.02 (98.69 to 123.36)	29.21 (13.81 to 44.62)	< 0.001	14	101.67 (89.33 to 114.00)	16.81 (1.41 to 32.22)	0.03
Elbow flexion								
Baseline	25	107.08 (98.90 to 115.26)	NA	NA	25	104.84 (96.66 to 113.02)	NA	NA
Week 1	23	111.48 (102.95 to 120.01)	4.40 (-7.42 to 16.22)	0.46	18	108.56 (98.91 to 118.20)	3.72 (-8.93 to 16.37)	0.56
Week 2	14	113.52 (102.59 to 124.46)	6.44 (-7.22 to 20.10)	0.35	14	112.00 (101.06 to 122.94)	7.16 (-6.50 to 20.82)	0.30

(continued)

Table 4. Continued.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	Mean (95% CI)	Change from baseline	p value	
Forearm pronation								
Baseline	25	47.20 (33.73 to 60.67)	NA	NA	25	50.88 (37.41 to 64.35)	NA	NA
Week 1	23	54.91 (40.87 to 68.95)	7.71 (-11.74 to 27.17)	0.43	18	52.43 (36.56 to 68.30)	1.55 (-19.27 to 22.36)	0.88
Week 2	14	47.48 (29.48 to 65.47)	0.28 (-22.20 to 22.75)	0.98	14	45.79 (27.79 to 63.78)	-5.10 (-27.57 to 17.38)	0.65
Forearm supination								
Baseline	25	48.79 (35.46 to 62.11)	NA	NA	25	51.83 (38.50 to 65.15)	NA	NA
Week 1	23	59.71 (45.82 to 73.60)	10.92 (-8.33 to 30.17)	0.26	18	51.76 (35.06 to 67.46)	-0.07 (-20.66 to 20.53)	0.995
Week 2	14	64.67 (46.86 to 82.47)	15.88 (-6.36 to 38.12)	0.16	14	51.07 (33.27 to 68.88)	-0.76 (-23.00 to 21.49)	0.95
Wrist flexion								
Baseline	25	20.75 (-39.74 to 81.24)	NA	NA	25	23.60 (-36.89 to 84.09)	NA	NA
Week 1	23	24.87 (-30.65 to 80.39)	4.12 (-8.79 to 17.03)	0.53	18	25.54 (-31.37 to 82.45)	1.94 (-11.88 to 15.75)	0.78
Week 2	14	25.45 (-27.78 to 78.69)	4.71 (-10.21 to 19.62)	0.53	14	31.50 (-21.74 to 84.74)	7.90 (-7.02 to 22.82)	0.30
Wrist extension								
Baseline	25	21.81 (12.67 to 30.96)	NA	NA	25	26.56 (17.42 to 35.71)	NA	NA
Week 1	23	28.35 (18.81 to 37.88)	6.53 (-6.68 to 19.75)	0.33	18	25.85 (15.07 to 36.63)	-0.71 (-14.84 to 13.43)	0.92
Week 2	14	24.00 (11.78 to 36.22)	2.19 (-13.08 to 17.45)	0.78	14	26.55 (14.33 to 38.77)	-0.01 (-15.28 to 15.25)	0.999

(continued)



Table 4. Continued.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	Mean (95% CI)	Change from baseline	p value	
Strength (N)								
Shoulder flexors								
Baseline	25	50.24 (18.86 to 81.63)	NA	NA	25	42.55 (11.16 to 73.93)	NA	NA
Week 1	22	61.03 (18.14 to 103.92)	10.79 (−6.43 to 28.01)	0.22	18	52.76 (−8.84 to 114.37)	10.22 (−7.99 to 28.43)	0.27
Week 2	14	64.08 (−2.36 to 130.52)	13.84 (−5.83 to 33.50)	0.17	14	54.28 (−12.16 to 120.71)	11.73 (−7.93 to 31.39)	0.24
Shoulder abductors								
Baseline	25	53.54 (42.49 to 64.59)	NA	NA	25	43.44 (32.39 to 54.49)	NA	NA
Week 1	22	62.27 (50.49 to 74.04)	8.72 (−7.43 to 24.87)	0.29	18	55.62 (42.60 to 68.64)	12.18 (−4.89 to 29.26)	0.16
Week 2	14	69.22 (54.45 to 83.98)	15.68 (−2.77 to 34.11)	0.10	14	57.59 (42.83 to 72.35)	14.15 (−4.29 to 32.59)	0.13
Elbow flexors								
Baseline	25	47.54 (35.69 to 59.39)	NA	NA	25	42.80 (30.95 to 54.66)	NA	NA
Week 1	22	60.02 (47.39 to 72.66)	12.48 (−4.84 to 29.81)	0.16	18	44.37 (30.40 to 58.34)	1.56 (−16.75 to 19.88)	0.87
Week 2	14	60.30 (44.46 to 76.14)	12.76 (−7.02 to 32.54)	0.20	14	48.89 (33.05 to 64.72)	6.08 (−13.70 to 25.86)	0.54
Wrist flexors								
Baseline	25	13.39 (−75.88 to 102.65)	NA	NA	25	12.37 (−76.89 to 101.64)	NA	NA
Week 1	22	15.06 (−32.25 to 62.38)	1.68 (−9.36 to 12.72)	0.76	18	16.26 (−27.58 to 60.10)	3.89 (−7.78 to 15.56)	0.51
Week 2	14	14.81 (−43.62 to 73.24)	1.43 (−11.18 to 14.03)	0.82	14	14.35 (−44.08 to 72.78)	1.98 (−10.62 to 14.58)	0.76

(continued)

Table 4. Continued.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	Mean (95% CI)	Change from baseline	p value	
Wrist extensors								
Baseline	25	21.67 (12.61 to 30.73)	NA	NA	25	15.34 (6.28 to 24.41)	NA	NA
Week 1	22	22.77 (13.11 to 32.43)	1.10 (-12.15 to 14.35)	0.87	18	18.44 (7.76 to 29.12)	3.09 (-10.91 to 17.10)	0.66
Week 2	14	18.67 (6.56 to 30.78)	-3.00 (-18.12 to 12.13)	0.70	14	19.10 (6.99 to 31.21)	3.75 (-11.37 to 18.88)	0.62
Arm and hand motor ability (Wolf Motor Function Test)								
Task completion time (seconds)								
Baseline	25	50.40 (-46.60 to 147.39)	NA	NA	25	52.37 (-44.63 to 149.36)	NA	NA
Week 1	19	49.65 (-47.64 to 146.95)	-0.74 (-21.82 to 20.34)	0.94	17	40.79 (-56.66 to 138.23)	-11.58 (-33.36 to 10.19)	0.29
Week 2	13	49.61 (-48.27 to 147.49)	-0.79 (-24.47 to 22.90)	0.95	14	43.92 (-53.83 to 141.67)	-8.45 (-31.57 to 14.67)	0.47
Quality of movement (score)								
Baseline	25	2.59 (1.82 to 3.36)	NA	NA	25	2.35 (1.58 to 3.12)	NA	NA
Week 1	19	2.54 (1.72 to 3.35)	-0.05 (-0.79 to 0.69)	0.90	17	2.80 (1.97 to 3.64)	0.45 (-0.31 to 1.22)	0.24
Week 2	13	2.63 (1.73 to 3.53)	0.04 (-0.79 to 0.88)	0.92	14	2.79 (1.90 to 3.67)	0.44 (-0.38 to 1.25)	0.29
Perceived upper limb motor function								
Baseline	25	3.72 (3.06 to 4.38)	NA	NA	25	4.00 (3.34 to 4.66)	NA	NA
Week 1	23	4.44 (3.74 to 5.13)	0.72 (-0.25 to 1.68)	0.14	18	5.33 (4.55 to 6.12)	1.33 (0.31 to 2.36)	0.01
Week 2	14	5.50 (4.61 to 6.39)	1.78 (0.67 to 2.89)	0.002	14	5.64 (4.76 to 6.53)	1.64 (0.53 to 2.75)	0.004

(continued)

Table 4. Continued.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	Mean (95% CI)	Change from baseline	p value	
Quality of life (EQ-5D-5L)								
Baseline	25	0.57 (0.50 to 0.65)	NA	NA	25	0.55 (0.47 to 0.63)	NA	NA
Week 1	23	0.66 (0.58 to 0.74)	0.09 (-0.02 to 0.20)	0.20	18	0.69 (0.60 to 0.77)	0.14 (0.02 to 0.25)	0.02
Week 2	14	0.70 (0.60 to 0.80)	0.13 (0.00 to 0.25)	0.05	14	0.73 (0.62 to 0.83)	0.18 (0.05 to 0.30)	0.01

CI, confidence interval; NA, not applicable.

upper limb strength, arm and hand motor ability (WMFT), perceived upper limb motor function and QoL (EQ-5D-5L) between the intervention and control groups, although significant improvements in coordination/speed (FMA-UE subscale), AROM of shoulder flexion and abduction, perceived upper limb motor function and QoL (EQ-5D-5L) over time were observed in either one or both of the groups. Per-protocol analyses revealed that the exercise system significantly improved the AROM of shoulder abduction at 2 weeks compared with the commercial games; significant improvements in coordination/speed (FMA-UE subscale), AROM of shoulder flexion, perceived upper limb motor function and QoL (EQ-5D-5L) over time were also observed in either one or both of the groups, but without significant between-group differences. In terms of safety, VR-induced side effects were mild and short lived, with no severe pain or significant increase in upper limb muscle spasticity observed, indicating that the exercise system was generally safe to use. Furthermore, the participants' overall perceptions of the exercise system were highly positive.

**Interpretation of study outcomes**

**Upper limb motor function (FMA-UE).** A possible explanation for failing to detect improvements in upper limb motor function could be because the VR exercise provided was insufficient. Although the ideal level of practice has yet to be determined,<sup>5,41</sup> some general guidelines on the amount of training have been proposed. For example, Laver et al.<sup>23</sup> reported that more than 15 hours of VR-based interventions tended to result in greater benefits in upper limb motor function than shorter interventions. Thus, a longer study period with a longer duration of training may provide different results than those we obtained in the present study. Moreover, in patients with mild impairments, the FMA-UE may not be responsive to changes (i.e. the ceiling effect).<sup>58</sup> In the present study, some of the participants were already in the Brunnstrom recovery stage 5 for the affected arm (performing complex movements is possible in this stage) at baseline, and thus, their FMA-UE scores were already close to the maximum possible score. Therefore, improvements in the upper limb motor function of these participants were limited.

**AROM.** The AROM results indicate that the exercise system has the potential to improve the shoulder joint movement range. This may be because the exercise system provided four types of exercises that focused separately on different joints, allowing the participants to learn to control a specific joint independently while inhibiting unwanted movements of other joints.<sup>59,60</sup>

**Strength.** A possible explanation for failing to detect improvements in muscle strength could be that most intervention participants' exercises focused on improving their

Table 5. Effects of the virtual reality-based exercise intervention on effectiveness outcomes based on the per-protocol analyses.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Mean (95% CI)	p value	n	Mean (95% CI)	Mean (95% CI)	
Upper limb motor function (Fugl-Meyer Assessment for Upper Extremity)								
Shoulder/elbow/forearm								
Baseline	14	16.07 (11.53 to 20.61)	NA	NA	14	19.43 (14.89 to 23.97)	NA	NA
Week 1	14	20.21 (15.67 to 24.76)	4.14 (-2.28 to 10.57)	0.20	14	22.57 (18.03 to 27.11)	3.14 (-3.28 to 9.57)	0.33
Week 2	14	21.86 (17.32 to 26.40)	5.79 (-0.64 to 12.21)	0.08	14	23.57 (19.03 to 28.11)	4.14 (-2.28 to 10.57)	0.20
Wrist								
Baseline	14	1.79 (-0.01 to 3.58)	NA	NA	14	2.43 (0.64 to 4.22)	NA	NA
Week 1	14	2.71 (0.92 to 4.51)	0.93 (-1.61 to 3.46)	0.47	14	3.71 (1.92 to 5.51)	1.29 (-1.25 to 3.82)	0.32
Week 2	14	3.43 (1.64 to 5.22)	1.64 (-0.89 to 4.18)	0.20	14	4.07 (2.28 to 5.87)	1.64 (-0.89 to 4.18)	0.20
Hand								
Baseline	14	5.86 (-8.65 to 20.37)	NA	NA	14	5.86 (-8.65 to 20.37)	NA	NA
Week 1	14	7.29 (-7.22 to 21.80)	1.43 (-2.19 to 5.04)	0.43	14	7.14 (-7.37 to 21.65)	1.29 (-2.33 to 4.90)	0.48
Week 2	14	7.43 (-7.08 to 21.94)	1.57 (-2.04 to 5.19)	0.39	14	8.36 (-6.15 to 22.87)	2.50 (-1.12 to 6.12)	0.17
Coordination/speed								
Baseline	14	4.21 (3.80 to 4.63)	NA	NA	14	4.21 (3.80 to 4.63)	NA	NA
Week 1	14	4.71 (4.30 to 5.13)	0.50 (-0.08 to 1.08)	0.09	14	4.71 (4.30 to 5.13)	0.50 (-0.08 to 1.08)	0.09
Week 2	14	4.86 (4.44 to 5.27)	0.64 (0.06 to 1.23)	0.03	14	4.71 (4.30 to 5.13)	0.50 (-0.08 to 1.08)	0.09

(continued)



Table 5. Continued.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	Mean (95% CI)	Change from baseline	p value	
Total								
Baseline	14	27.93 (19.51 to 36.35)	NA	NA	14	31.93 (23.51 to 40.35)	NA	NA
Week 1	14	34.93 (26.51 to 43.35)	7.00 (-4.90 to 18.90)	0.25	14	38.14 (29.73 to 46.56)	6.21 (-5.69 to 18.12)	0.30
Week 2	14	37.57 (29.15 to 45.99)	9.64 (-2.26 to 21.55)	0.11	14	40.71 (32.30 to 49.13)	8.79 (-3.12 to 20.69)	0.15
Active range of motion (°)								
Shoulder flexion								
Baseline	14	89.95 (79.02 to 100.88)	NA	NA	14	101.29 (90.36 to 112.22)	NA	NA
Week 1	14	106.98 (96.05 to 117.91)	17.03 (1.57 to 32.48)	0.03	14	107.57 (96.64 to 118.50)	6.29 (-9.17 to 21.74)	0.42
Week 2	14	118.31 (107.38 to 129.24)	28.36 (12.90 to 43.82)	< 0.001	14	113.81 (102.88 to 124.71)	12.52 (-2.94 to 27.98)	0.11
Shoulder abduction								
Baseline	14	74.79 (63.67 to 85.90)	NA	NA	14	88.71 (77.60 to 99.83)	NA	NA
Week 1	14	92.36 (81.25 to 103.47)	17.57 (1.86 to 33.29)	0.03	14	97.12 (86.01 to 108.23)	8.41 (-7.31 to 24.12)	0.29
Week 2	14	111.02 (99.91 to 122.14)	36.24 (20.53 to 51.95)	< 0.001	14	101.67 (90.56 to 112.78)	12.95 (-2.76 to 28.67)	0.11
Elbow flexion								
Baseline	14	102.00 (49.13 to 154.87)	NA	NA	14	99.57 (46.70 to 152.45)	NA	NA
Week 1	14	110.12 (57.25 to 162.99)	8.12 (-6.58 to 22.82)	0.28	14	107.83 (54.96 to 160.71)	8.26 (-6.44 to 22.96)	0.27
Week 2	14	113.52 (60.65 to 166.40)	11.52 (-3.18 to 26.22)	0.12	14	112.00 (59.13 to 164.87)	12.43 (-2.27 to 27.13)	0.10

(continued)

Table 5. Continued.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	Mean (95% CI)	Change from baseline	p value	
Forearm pronation								
Baseline	14	36.00 (17.38 to 54.62)	NA	NA	14	44.19 (25.57 to 62.81)	NA	NA
Week 1	14	44.67 (26.05 to 63.29)	8.67 (-17.66 to 35.00)	0.51	14	49.22 (30.60 to 67.83)	5.02 (-21.31 to 31.36)	0.71
Week 2	14	47.48 (28.86 to 66.09)	11.48 (-14.86 to 37.81)	0.39	14	45.79 (27.17 to 64.40)	1.60 (-24.74 to 27.93)	0.90
Forearm supination								
Baseline	14	43.45 (-197.02 to 283.92)	NA	NA	14	42.81 (-197.66 to 283.28)	NA	NA
Week 1	14	55.22 (-185.26 to 295.69)	11.76 (-13.61 to 37.14)	0.36	14	48.48 (-191.99 to 288.95)	5.67 (-19.71 to 31.04)	0.66
Week 2	14	64.67 (-175.80 to 305.14)	21.21 (-4.16 to 46.59)	0.10	14	51.07 (-189.40 to 291.54)	8.26 (-17.11 to 33.63)	0.52
Wrist flexion								
Baseline	14	14.36 (2.55 to 26.17)	NA	NA	14	42.81 (-197.66 to 283.28)	NA	NA
Week 1	14	22.22 (10.40 to 34.03)	7.86 (-8.85 to 24.56)	0.35	14	48.48 (-191.99 to 288.95)	5.17 (-11.54 to 21.87)	0.54
Week 2	14	25.45 (13.64 to 37.27)	11.10 (-5.61 to 27.80)	0.19	14	51.07 (-189.40 to 291.54)	14.33 (-2.37 to 31.04)	0.09
Wrist extension								
Baseline	14	14.36 (2.55 to 26.17)	NA	NA	14	17.17 (5.35 to 28.98)	NA	NA
Week 1	14	22.22 (10.40 to 34.03)	7.12 (-10.63 to 24.86)	0.43	14	22.33 (10.52 to 34.15)	0.43 (-17.32 to 18.17)	0.96
Week 2	14	25.45 (13.64 to 37.27)	8.31 (-9.43 to 26.06)	0.35	14	31.50 (19.69 to 43.31)	4.19 (-13.55 to 21.94)	0.64

(continued)

Table 5. Continued.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	Mean (95% CI)	Change from baseline	p value	
Strength (N)								
Shoulder flexors								
Baseline	14	46.89 (31.99 to 61.78)	NA	NA	44.52 (29.63 to 59.42)	NA	NA	NA
Week 1	14	55.32 (40.43 to 70.22)	8.44 (-12.43 to 29.30)	0.42	48.60 (33.71 to 63.50)	4.08 (-16.78 to 24.94)	0.70	0.77
Week 2	14	64.08 (49.19 to 78.97)	17.19 (-3.67 to 38.05)	0.11	54.28 (39.38 to 69.17)	9.75 (-11.11 to 30.61)	0.36	0.62
Shoulder abductors								
Baseline	14	52.07 (-9.62 to 113.76)	NA	NA	42.66 (-19.03 to 104.35)	NA	NA	NA
Week 1	14	58.11 (-3.58 to 119.80)	6.04 (-14.26 to 26.33)	0.56	51.49 (-10.20 to 113.18)	8.83 (-11.46 to 29.12)	0.39	0.85
Week 2	14	69.22 (7.53 to 130.90)	17.14 (-3.15 to 37.44)	0.10	57.59 (-4.10 to 119.28)	14.93 (-5.36 to 35.23)	0.15	0.88
Elbow flexors								
Baseline	14	42.72 (29.03 to 56.40)	NA	NA	40.37 (26.69 to 54.06)	NA	NA	NA
Week 1	14	56.14 (42.45 to 69.82)	13.42 (-5.93 to 32.78)	0.17	42.19 (28.51 to 55.88)	1.82 (-17.54 to 21.17)	0.85	0.40
Week 2	14	60.30 (46.62 to 73.99)	17.59 (-1.77 to 36.94)	0.07	48.89 (35.20 to 62.57)	8.51 (-10.84 to 27.87)	0.38	0.51
Wrist flexors								
Baseline	14	8.85 (-0.81 to 18.51)	NA	NA	9.81 (0.15 to 19.47)	NA	NA	NA
Week 1	14	9.80 (0.14 to 19.46)	0.95 (-12.71 to 14.61)	0.89	12.91 (3.26 to 22.57)	3.11 (-10.56 to 16.77)	0.65	0.83
Week 2	14	14.81 (5.15 to 24.47)	5.96 (-7.70 to 19.62)	0.39	14.35 (4.69 to 24.01)	4.54 (-9.12 to 18.20)	0.51	0.88

(continued)

Table 5. Continued.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	Mean (95% CI)	Change from baseline	p value	
<b>Wrist extensors</b>								
Baseline	14	11.96 (-63.77 to 87.69)	NA	NA	13.29 (-62.45 to 89.02)	NA	NA	NA
Week 1	14	13.80 (-61.94 to 89.53)	1.84 (-12.94 to 16.61)	0.81	14.64 (-61.10 to 90.37)	1.35 (-13.42 to 16.12)	0.86	0.96
Week 2	14	18.67 (-57.06 to 94.41)	6.71 (-8.06 to 21.49)	0.37	19.10 (-56.64 to 94.83)	5.81 (-8.96 to 20.58)	0.44	0.93
<b>Arm and hand motor ability (Wolf Motor Function Test)</b>								
<b>Task completion time (seconds)</b>								
Baseline	13	57.63 (35.96 to 79.30)	NA	NA	57.19 (36.13 to 78.25)	NA	NA	NA
Week 1	13	50.75 (29.08 to 72.42)	-6.88 (-34.08 to 20.32)	0.62	48.09 (27.03 to 69.14)	-9.10 (-35.32 to 17.11)	0.49	0.91
Week 2	13	49.61 (27.93 to 71.28)	-8.02 (-35.23 to 19.18)	0.56	43.92 (22.86 to 64.97)	-13.27 (-39.48 to 12.94)	0.32	0.78
<b>Quality of movement (score)</b>								
Baseline	13	2.26 (0.75 to 3.77)	NA	NA	2.12 (0.62 to 3.62)	NA	NA	NA
Week 1	13	2.48 (0.97 to 3.99)	0.23 (-0.72 to 1.17)	0.64	2.56 (1.06 to 4.06)	0.43 (-0.48 to 1.34)	0.35	0.75
Week 2	13	2.63 (1.12 to 4.14)	0.37 (-0.57 to 1.32)	0.43	2.79 (1.29 to 4.29)	0.66 (-0.25 to 1.57)	0.15	0.66
<b>Perceived upper limb motor function</b>								
Baseline	14	3.36 (2.48 to 4.24)	NA	NA	4.36 (3.48 to 5.24)	NA	NA	NA
Week 1	14	4.43 (3.55 to 5.31)	1.07 (-0.18 to 2.32)	0.09	5.43 (4.55 to 6.31)	1.07 (-0.18 to 2.32)	0.09	> 0.99
Week 2	14	5.50 (4.62 to 6.38)	2.14 (0.90 to 3.39)	0.001	5.64 (4.76 to 6.52)	1.29 (0.04 to 2.53)	0.04	0.34

(continued)

Table 5. Continued.

Outcome	n	Intervention			Control			p value for between-group difference in mean changes
		Mean (95% CI)	Change from baseline	p value	Mean (95% CI)	Change from baseline	p value	
Quality of life (EQ-5D-5L)								
Baseline	14	0.60 (0.51 to 0.69)	NA	NA	14	0.57 (0.48 to 0.67)	NA	NA
Week 1	14	0.67 (0.57 to 0.76)	0.07 (-0.07 to 0.20)	0.32	14	0.70 (0.61 to 0.79)	0.13 (-0.004 to 0.26)	0.52
Week 2	14	0.70 (0.61 to 0.79)	0.10 (-0.03 to 0.23)	0.14	14	0.73 (0.63 to 0.82)	0.15 (0.02 to 0.28)	0.57

CI, confidence interval; NA, not applicable.

joint motion and facilitating normal movement patterns. Over 50% of the intervention participants (Table 2) were in Brunnstrom recovery stage 3 for the affected arm at baseline. For patients in this stage, performing movements with undue effort could lead to an abnormal increase in muscle spasticity, which would reinforce abnormal movement patterns and cause pain.<sup>61,62</sup> The controller used in the study weighed about 0.4 kg; thus, we were quite prudent to add additional weights (i.e. put weight on the wrist or hand) for those participants, to avoid adverse training effects. For the participants who were cleared for adding additional weights during exercises, approximately 35 minutes of exercises within 2 weeks might have been insufficient to significantly enhance muscle strength.<sup>63</sup> A review by Harris and Eng<sup>64</sup> indicated that, on average, strength training should be administered for at least 4 weeks (60 minutes/day, 2–3 days/week) to achieve promising improvements in upper limb muscle strength in poststroke patients. American Heart Association/American Stroke Association suggested that strength training should be performed 2 to 3 days per week, including 1 to 3 sets of 8 to 10 different exercises with 10 to 15 repetitions.<sup>65</sup>

**Arm and hand motor ability (WMFT).** A potential reason for failing to detect a significant improvement in arm and hand motor ability is that half of the tasks in the WMFT require the use of the hand and fingers, such as grasping a pencil and flipping cards over. In this study, the participants needed to just hold and move the controller when interacting with virtual environments, which scarcely involved any hand and finger movements; thus, training in hand and finger movements was limited.

**Perceived upper limb motor function.** The participants in both the groups perceived that their upper limb motor function significantly improved, and no between-group differences were observed. For stroke patients, being afflicted with long-term motor impairment can lead to low self-efficacy and negative health behaviours (e.g., decreased independence and autonomy in disease management).<sup>66</sup> Improving self-efficacy is essential for effective disease self-management.<sup>66</sup> Our results suggest that, combined with conventional rehabilitation therapy, either the exercise system or commercial games could positively impact participants' perceptions of their motor function, which may increase their self-efficacy in performing therapeutic exercises and recovering from motor impairment.

**QoL (EQ-5D-5L).** The control group showed significantly improved QoL, but this improvement was not significantly different from that in the intervention group. According to the EQ-5D-5L questionnaire, QoL can be assessed in multiple dimensions, including in mobility, self-care, usual activities, pain/discomfort and anxiety/depression dimensions. In this study, a higher proportion of control

**Table 6.** Distribution of the participants' ratings for pain in the affected upper limb due to the use of the immersive virtual reality-based exercise system.

Pain level	Intervention		Control	
	Week 1 (n = 23)	Week 2 (n = 14)	Week 1 (n = 18)	Week 2 (n = 14)
0	15 (65.2%)	10 (71.4%)	15 (83.3%)	11 (78.6%)
1	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (7.1%)
2	3 (13.0%)	1 (7.1%)	2 (11.1%)	1 (7.1%)
3	0 (0.0%)	0 (0.0%)	1 (5.6%)	0 (0.0%)
4	3 (13.0%)	3 (21.4%)	0 (0.0%)	1 (7.1%)
5	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
6	2 (8.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
7	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
8	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
9	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
10	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

participants reported no problems or slight problems in mobility after 1 and 2 weeks (Appendix 2). This might be the possible explanation for the results.

**Safety.** The safety of immersive VR-based interventions for poststroke motor rehabilitation has rarely been reported. In this study, immersive VR induced discomfort, such as dizziness and eye strain, but it was mild and short lived, indicating that immersive VR for poststroke exercise was safe overall for and well tolerated by the participants.

**Patients' perceptions.** The results indicate that the participants perceived that using the exercise system for upper limb exercises was easy, useful, enjoyable and satisfactory, which could have positively influenced their attitudes toward and intention to use the exercise system.<sup>28–31,36,37,54,55,67</sup> These findings reflect that the participants' overall perceptions of the exercise system were positive. However, these highly positive perceptions may be subject to bias, because the participants who had positive perceptions were more likely to have adhered to the intervention than those who had negative perceptions, who might have been those who withdrew from the study. It is suggested that the features of the exercise system that positively and negatively influence

patients' perceptions of the exercise system should be further identified, as this will guide improvements in the design of such interventions and thereby enhance poststroke patients' experience with and acceptance of such interventions.

### Implications for research

The study's findings have several implications for research. First, studies with larger sample sizes than this study are required to provide more robust evidence on the use of immersive VR-based interventions for poststroke upper limb motor rehabilitation. Based on the mean changes in the FMA-UE total score at week 2 (Table 5) in the two groups, to detect a mean difference of 6 points (the minimal clinical important change<sup>68,69</sup>) in the FMA-UE total score between groups with a SD value of 16 (calculated based on results in Table 5), assuming a power of 80% and attrition rate of 20%, the total sample size should be 140.<sup>70–72</sup> Stroke patients with different clinical characteristics should also be included and stratified in future studies to examine how different groups of patients respond to such interventions. Second, the exercise system did not produce clinically significant improvements in study outcomes during our 2-week study period, but its long-term effects remain unknown. A longer intervention duration is needed to provide evidence regarding the long-term effects of such a system on outcomes of poststroke upper limb motor rehabilitation. Third, the games used in this study did not involve all gross and fine movements of the upper limb joints. Additionally, the games mainly focused on single-joint movements, and multi-joint movements were less involved (only the reaching exercise). Thus, it is recommended that VR-based games that include more types of movements of upper limb joints be designed and developed for poststroke upper limb motor rehabilitation. For example, sensor-based hand tracking systems, such as Ultraleap 3Di, can be combined with HMD devices to provide exercises for hand and finger movements. Finally, possible reasons for participant withdrawal should be examined in future studies, as these may shed light on factors that impede patients' acceptance of and adherence to VR-based interventions. Such information would be useful for improving the design, development and implementation of VR-based interventions, to improve patients' experiences with and acceptance of such interventions and increase their benefit for patients. In general, challenges that are frequently seen in healthcare field research should be carefully considered to increase their effectiveness.<sup>73</sup>

### Implications for practice

The study offers implications for practice. First, our findings suggest that the exercise system is useful for supporting the delivery of upper limb exercises for poststroke patients, given that the intervention was regarded as tolerable and satisfactory overall by the participants and



**Table 7.** Distribution of the participants' muscle spasticity in the affected upper limb.

Muscle spasticity level	Intervention		Control	
	Week 1 (n = 23)	Week 2 (n = 14)	Week 1 (n = 18)	Week 2 (n = 14)
<b>Biceps</b>				
0	19 (82.6%)	9 (64.3%)	15 (83.3%)	12 (85.7%)
1	3 (13.0%)	3 (21.4%)	2 (11.1%)	1 (7.1%)
1+	1 (4.3%)	2 (14.3%)	1 (5.6%)	1 (7.1%)
2	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
3	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
4	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
<b>Triceps</b>				
0	22 (95.7%)	13 (92.9%)	17 (94.4%)	12 (85.7%)
1	1 (4.3%)	1 (7.1%)	1 (5.6%)	2 (14.3%)
1+	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
2	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
3	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
4	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
<b>Forearm pronators</b>				
0	23 (100.0%)	13 (92.9%)	18 (100.0%)	14 (100.0%)
1	0 (0.0%)	1 (7.1%)	0 (0.0%)	0 (0.0%)
1+	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
2	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
3	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
4	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
<b>Forearm supinators</b>				
0	21 (91.3%)	12 (85.7%)	15 (83.3%)	12 (85.7%)
1	2 (8.7%)	2 (14.3%)	3 (16.7%)	2 (14.3%)
1+	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
2	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

(continued)

Table 7. Continued.

Muscle spasticity level	Intervention		Control	
	Week 1 (n = 23)	Week 2 (n = 14)	Week 1 (n = 18)	Week 2 (n = 14)
3	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
4	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Wrist flexors				
0	23 (100.0%)	14 (100.0%)	18 (100.0%)	14 (100.0%)
1	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
1+	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
2	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
3	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
4	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Wrist extensors				
0	18 (78.3%)	10 (71.4%)	16 (88.9%)	10 (71.4%)
1	5 (21.7%)	1 (7.1%)	2 (11.1%)	3 (21.4%)
1+	0 (0.0%)	2 (14.3%)	0 (0.0%)	1 (7.1%)
2	0 (0.0%)	1 (7.1%)	0 (0.0%)	0 (0.0%)
3	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
4	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

demonstrated the potential to improve the upper limb joints movement range. Second, patients may experience discomfort (e.g. dizziness) with immersive VR devices, and thus, clinicians and therapists should carefully monitor and identify susceptible patients when implementing VR-based therapy.<sup>74,75</sup> Extra care should be taken, such as by limiting the duration of exposure to virtual environments, informing patients about appropriate strategies for interacting with virtual environments (e.g., avoiding quick head movements) and monitoring patients to identify possible signs of discomfort.<sup>75</sup> In addition, elements that may induce discomfort, such as poor environmental illumination, poor contrast and highly realistic graphics, should be avoided in the design of virtual environments.<sup>75-78</sup>

### Limitations

This study has the following limitations. First, the study had a small and heterogenous sample and thus was only a feasibility and proof-of-concept study. It was also less likely to

perform sufficiently powered subgroup analysis using potentially important patient clinical characteristics (e.g. time from stroke onset and the severity of motor impairment). Second, due to the local medical policy about hospitalization length, the study period was short. Consequently, the total duration of the VR-based exercise intervention provided for the participants was less than the suggested total duration of training needed to produce evident improvements in upper limb motor function (i.e., more than 15 hours of total intervention).<sup>23</sup> Therefore, the effects of the exercise system might have been underestimated. Third, the study had a high attrition rate. Some withdrawals might have been due to dissatisfaction with the exercise system. Therefore, the participants' perceptions might have been overestimated. In addition, the participants' reasons for withdrawal were not collected, so their barriers to accepting and adhering to the intervention and the usability problems of the system remain unknown.

## Conclusions

This proof-of-concept RCT was conducted to develop, implement and evaluate an immersive VR-based exercise system for supporting upper limb exercises in poststroke patients. This exercise system holds promise for improving the shoulder joint range of motion and was found to be safe to use. Furthermore, the participants were overall positive about using the exercise system for upper limb exercises. This RCT demonstrates that using immersive VR for post-stroke motor rehabilitation is potentially effective, safe and acceptable in poststroke patients. These findings can serve as a basis for larger-scale RCTs of immersive VR-based systems for poststroke upper limb motor rehabilitation.

**Acknowledgements:** The authors acknowledge the support from Dingzhou People's Hospital and The University of Hong Kong—Shenzhen Hospital. The authors are grateful for the technical assistance received from Mr Jason Wong, Ms. Aki Dai, Mr James Chan, Dr Henry Lau, Dr Leith Chan, Dr Tobey Ko and Mr Charles Chan. The authors also acknowledge the clinical advice and assistance for participant recruitment and randomization received from Dr Yali Zhao, Dr Minzhen Zhen, Dr Bohua Zhang, Dr Hongyu Kang, Dr Jing Li, Dr Nan Li (F), Dr Nan Li (M), Dr Bin Meng, Dr Ling Tong, Dr Pei Yang, Dr Yun Zhang, Ms. Danhui Zhong, Ms. Hailan Li, Ms. Rong He and Ms. Tianrong Chen.

**Contributorship:** JC and CKLO designed the study. ZL and EHKY supported participant recruitment and experiment implementation. JC, YZ and TH conducted data collection and outcome assessment. JC performed data analysis and drafted the manuscript. CKLO reviewed and significantly revised the manuscript.

**Declaration of conflicting interests:** The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

**Funding:** The authors disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was supported by the Seed Funding for Basic Research, The University of Hong Kong (grant number 201811159012).

**Copyright:** The authors declare that they have obtained the permission to use the figures in this study.

**Guarantor:** CKLO.

**Ethical approval:** The trial was approved by the Human Research Ethics Committee of The University of Hong Kong-Shenzhen Hospital (approval number: [2021]137) and Dingzhou People's Hospital (approval number: hx-2021-04) in China and has been registered in the Chinese Clinical Trial Registry (registration number: ChiCTR2100047150). All participants provided written informed consent before participating in the study.

**ORCID iD:** Calvin Kalun Or  <https://orcid.org/0000-0002-9819-8865>

## References

- Cheung KL, et al. Neuroplasticity and virtual reality. In: Weiss PL, Keshner EA and Levin MF (eds) *Virtual reality for physical and motor rehabilitation*. New York: Springer, 2014, pp.5–24.
- Lewis GN, et al. Virtual reality games for rehabilitation of people with stroke: perspectives from the users. *Disability and Rehabilitation: Assistive Technology* 2011; 6: 453–463.
- Lohse K, et al. Video games and rehabilitation: using design principles to enhance engagement in physical therapy. *J Neurol Phys Ther* 2013; 37: 166–175.
- Levac DE and Sveistrup H. Motor learning and virtual reality. In: Weiss PL, Keshner EA and Levin MF (eds) *Virtual reality for physical and motor rehabilitation*. New York: Springer, 2014, pp.25–46.
- Levin MF and Demers M. Motor learning in neurological rehabilitation. *Disabil Rehabil* 2021; 43: 3445–3453.
- Chen J, Xie Z and Or C. Effectiveness of immersive virtual reality-supported interventions for patients with disorders or impairments: a systematic review and meta-analysis. *Health Technol (Berl)* 2021; 11: 811–833.
- Islam MS and Lim S. Vibrotactile feedback in virtual motor learning: a systematic review. *Appl Ergon* 2022; 101: 103694.
- Baragash RS, Aldowah H and Ghazal S. Virtual and augmented reality applications to improve older adults' quality of life: a systematic mapping review and future directions. *Digital Health* 2022; 8: 20552076221132099.
- Dhar E, et al. A scoping review to assess the effects of virtual reality in medical education and clinical care. *Digital Health* 2023; 9: 20552076231158022.
- Fusco A and Tieri G. Challenges and perspectives for clinical applications of immersive and non-immersive virtual reality. *J Clin Med* 2022; 11: 4540.
- Or CK, Duffy VG and Cheung CC. Perception of safe robot idle time in virtual reality and real industrial environments. *Int J Ind Ergon* 2009; 39: 807–812.
- Duffy VG, Or CK and Lau VW. Perception of safe robot speed in virtual and real industrial environments. *Human Factors and Ergonomics in Manufacturing & Service Industries* 2006; 16: 369–383.
- Helou S, et al. Virtual reality for healthcare: a scoping review of commercially available applications for head-mounted displays. *Digital Health* 2023; 9: 20552076231178619.
- Pallavicini F, Pepe A and Minissi ME. Gaming in virtual reality: what changes in terms of usability, emotional response and sense of presence compared to non-immersive video games? *Simul Gaming* 2019; 50: 136–159.
- Sanchez-Vives MV and Slater M. From presence to consciousness through virtual reality. *Nat Rev Neurosci* 2005; 6: 332–339.
- Chen J and Or C. Assessing the use of immersive virtual reality, mouse and touchscreen in pointing and dragging-and-dropping tasks among young, middle-aged and older adults. *Appl Ergon* 2017; 65: 437–448.
- Vecchiato G, et al. Electroencephalographic correlates of sensorimotor integration and embodiment during the appreciation

- of virtual architectural environments. *Front Psychol* 2015; 6: 1944.
18. Iosa M, et al. The Michelangelo Effect: art improves the performance in a virtual reality task developed for upper limb neurorehabilitation. *Front Psychol* 2021; 11: 611956.
  19. Chen J, Xie Z and Or CK. Head-mounted display virtual reality in disease treatment: a systematic review and meta-analysis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 2020; 64: 1388–1389.
  20. Chen J, Or CK and Chen T. Effectiveness of using virtual reality-supported exercise therapy for upper extremity motor rehabilitation in patients with stroke: systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res* 2022; 24: e24111.
  21. Domínguez-Téllez P, et al. Game-based virtual reality interventions to improve upper limb motor function and quality of life after stroke: systematic review and meta-analysis. *Games Health J* 2020; 9: 1–10.
  22. Saposnik G and Levin M. Virtual reality in stroke rehabilitation: a meta-analysis and implications for clinicians. *Stroke* 2011; 42: 1380–1386.
  23. Laver KE, et al. Virtual reality for stroke rehabilitation. *Cochrane Database Syst Rev* 2017; CD008349.
  24. Mekbib DB, et al. Virtual reality therapy for upper limb rehabilitation in patients with stroke: a meta-analysis of randomized clinical trials. *Brain Inj* 2020; 34: 456–465.
  25. Aminov A, et al. What do randomized controlled trials say about virtual rehabilitation in stroke? A systematic literature review and meta-analysis of upper-limb and cognitive outcomes. *J Neuroeng Rehabil* 2018; 15: 29.
  26. Lohse KR, et al. Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. *PLOS One* 2014; 9: e93318.
  27. Or C and Tao D. Usability study of a computer-based self-management system for older adults with chronic diseases. *JMIR Res Protoc* 2012; 1: e13.
  28. Or CKL and Karsh BT. A systematic review of patient acceptance of consumer health information technology. *J Am Med Inform Assoc* 2009; 16: 550–560.
  29. Or CKL, et al. Factors affecting home care patients' acceptance of a web-based interactive self-management technology. *J Am Med Inform Assoc* 2011; 18: 51–59.
  30. Chen T, et al. Path analysis of the roles of age, self-efficacy, and TAM constructs in the acceptance of performing upper limb exercises through immersive virtual reality games. *Int J Ind Ergon* 2022; 91: 103360.
  31. Yan M and Or C. A 12-week pilot study of acceptance of a computer-based chronic disease self-monitoring system among patients with type 2 diabetes mellitus and/or hypertension. *Health Informatics J* 2019; 25: 828–843.
  32. Xie Z, et al. An examination of the socio-demographic correlates of patient adherence to self-management behaviors and the mediating roles of health attitudes and self-efficacy among patients with coexisting type 2 diabetes and hypertension. *BMC Public Health* 2020; 20: 1227.
  33. Liu K, et al. A longitudinal examination of tablet self-management technology acceptance by patients with chronic diseases: integrating perceived hand function, perceived visual function, and perceived home space adequacy with the TAM and TPB. *Appl Ergon* 2022; 100: 103667.
  34. Karsh BT, Holden RJ and Or CKL. Human factors and ergonomics of health information technology implementation. In: Carayon P (eds) *Handbook of human factors and ergonomics in health care and patient safety*. Boca Raton: CRC Press, 2011, pp.249–264.
  35. Or CK, Holden RJ, Valdez RS, et al. Human factors engineering and user-centered design for mobile health technology: enhancing effectiveness, efficiency, and satisfaction. In: Duffy VG (eds) *Human-Automation interaction: mobile computing*. Cham.: Springer International Publishing, 2023, pp.97–118.
  36. Davis FD, Bagozzi RP and Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. *Manage Sci* 1989; 35: 982–1003.
  37. Venkatesh V, et al. User acceptance of information technology: toward a unified view. *MIS Q* 2003; 27: 425–478.
  38. Toh SFM, Gonzalez PC and Fong KN. Usability of a wearable device for home-based upper limb telerehabilitation in persons with stroke: a mixed-methods study. *Digital Health* 2023; 9: 20552076231153737.
  39. Chen T and Or CK. Perceptions of a machine learning-based lower-limb exercise training system among older adults with knee pain. *Digital Health* 2023; 9: 20552076231186069.
  40. Kwakkel G, Kollen B and Lindeman E. Understanding the pattern of functional recovery after stroke: facts and theories. *Restor Neurol Neurosci* 2004; 22: 281–299.
  41. Langhorne P, Bernhardt J and Kwakkel G. Stroke rehabilitation. *Lancet* 2011; 377: 1693–1702.
  42. Harris JE, et al. A self-administered Graded Repetitive Arm Supplementary Program (GRASP) improves arm function during inpatient stroke rehabilitation: a multi-site randomized controlled trial. *Stroke* 2009; 40: 2123–2128.
  43. Hebert D, et al. Canadian stroke best practice recommendations: stroke rehabilitation practice guidelines, update 2015. *Int J Stroke* 2016; 11: 459–484.
  44. Pang MY, Harris JE and Eng JJ. A community-based group upper extremity exercise program improves motor function and performance of functional activities in chronic stroke: a randomized controlled trial. *Arch Phys Med Rehabil* 2006; 87: 1–9.
  45. Fugl-Meyer AR, et al. The post-stroke hemiplegic patient. I: a method for evaluation of physical performance. *Scand J Rehabil Med* 1975; 7: 13–31.
  46. Gajdosik RL and Bohannon RW. Clinical measurement of range of motion: review of goniometry emphasizing reliability and validity. *Phys Ther* 1987; 67: 1867–1872.
  47. Norikin CC and White DJ. *Measurement of joint motion: a guide to goniometry*. 5th ed. FA Davis, 2016.
  48. Bohannon RW. Test-retest reliability of hand-held dynamometry during a single session of strength assessment. *Phys Ther* 1986; 66: 206–209.
  49. Bohannon RW and Andrews AW. Interrater reliability of hand-held dynamometry. *Phys Ther* 1987; 67: 931–933.
  50. Mentiplay BF, et al. Assessment of lower limb muscle strength and power using hand-held and fixed dynamometry: a reliability and validity study. *PLOS One* 2015; 10: e0140822.
  51. Stark T, et al. Hand-held dynamometry correlation with the gold standard isokinetic dynamometry: a systematic review. *Physical Medicine and Rehabilitation* 2011; 3: 472–479.

52. Wolf SL, et al. Assessing Wolf Motor Function Test as outcome measure for research in patients after stroke. *Stroke* 2001; 32: 1635–1639.
53. Luo N, et al. Estimating an EQ-5D-5L value set for China. *Value Health* 2017; 20: 662–669.
54. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction* 1995; 7: 57–78.
55. Taylor S and Todd PA. Understanding information technology usage: a test of competing models. *Inf Syst Res* 1995; 6: 144–176.
56. West BT. Analyzing longitudinal data with the linear mixed models procedure in SPSS. *Eval Health Prof* 2009; 32: 207–228.
57. Jakobsen JC, et al. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Med Res Methodol* 2017; 17: 162.
58. Gladstone DJ, Danells CJ and Black SE. The Fugl-Meyer Assessment of motor recovery after stroke: a critical review of its measurement properties. *Neurorehabil Neural Repair* 2002; 16: 232–240.
59. Brunnstrom S. *Movement therapy in hemiplegia: a neurophysiological approach*. 1st ed. New York: Harper & Row, 1970.
60. Twichell TE. The restoration of motor function following hemiplegia in man. *Brain* 1951; 74: 443–480.
61. Brunnstrom S. Motor testing procedures in hemiplegia: based on sequential recovery stages. *Phys Ther* 1966; 46: 357–375.
62. Naghdi S, et al. A neurophysiological and clinical study of Brunnstrom recovery stages in the upper limb following stroke. *Brain Inj* 2010; 24: 1372–1378.
63. Salter K, Musovic A and Taylor NF. In the first 3 months after stroke is progressive resistance training safe and does it improve activity? A systematic review. *Top Stroke Rehabil* 2016; 23: 366–375.
64. Harris JE and Eng JJ. Strength training improves upper-limb function in individuals with stroke: a meta-analysis. *Stroke* 2010; 41: 136–140.
65. Billinger SA, et al. Physical activity and exercise recommendations for stroke survivors: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2014; 45: 2532–2553.
66. Bedrov A and Bulaj G. Improving self-esteem with motivational quotes: opportunities for digital health technologies for people with chronic disorders. *Front Psychol* 2018; 9: 2126.
67. Venkatesh V, Thong JYL and Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Q* 2012; 36: 157–178.
68. Arya KN, Verma R and Garg RK. Estimating the minimal clinically important difference of an upper extremity recovery measure in subacute stroke patients. *Top Stroke Rehabil* 2011; 18: 599–610.
69. Feys HM, et al. Effect of a therapeutic intervention for the hemiplegic upper limb in the acute phase after stroke : a single-blind, randomized, controlled multicenter trial. *Stroke* 1998; 29: 785–792.
70. Wittes J. Sample size calculations for randomized controlled trials. *Epidemiol Rev* 2002; 24: 39–53.
71. Zhong B. How to calculate sample size in randomized controlled trial? *J Thorac Dis* 2009; 1: 51–54.
72. Higgins JPT, et al. *Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022)*. 2022, Cochrane.
73. Holden RJ, et al. Field-based human factors in home and community settings: challenges and strategies. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 2015; 59: 562–566.
74. Nichols S, et al. *Incidence of virtual reality induced symptoms and effects (VRISE) in desktop and projection screen displays systems*, HSE Contract Research Report, Editor. 2000, Health and Safety Executive.
75. Sharples S, et al. Virtual reality induced symptoms and effects (VRISE): comparison of head mounted display (HMD), desktop and projection display systems. *Displays* 2008; 29: 58–69.
76. Chattha UA, et al. Motion sickness in virtual reality: an empirical evaluation. *IEEE Access* 2020; 8: 130486–130499.
77. Davis S, Nesbitt K and Nalivaiko E. Comparing the onset of cybersickness using the Oculus Rift and two virtual roller coasters. In: *Proceedings of 11th Australasian Conference on Interactive Entertainment*, Sydney, Australia, 2015.
78. Kobayashi N, et al. Effects of illuminance environment on visual induced motion sickness. In: *IEEE 7th Global Conference on Consumer Electronics*, Nara, Japan, 2018.

## Appendix 1. Commercial games used in the control group.

**Chocolat Rush** Chocolat Rush was a shooting game. The participants stood in the centre of a room in the virtual environment. The participants needed to hold the controller (presented as a gun in VR) and press a trigger button on the controller using their index finger to fire bullets at the monsters swarming in from different directions and blow them up. The participants achieved higher scores with more monsters hit. This game had no time limit.

**Zooma** The participants stood on a high platform in a hall in the virtual environment. A string of balls of different colours was rolling in front of the participants. The participants needed to hold the controller (presented as a gun in VR) and shoot different colours of balls. The participants needed to eliminate all the rolling balls by lining up three or more balls of the same colour. The participants achieved

higher scores with more balls eliminated. Each trial lasted approximately 2 minutes.

**Fruit Ninja VR** Fruits popped up from the ground in front of the participants in the virtual environment. The participants needed to hold the controller (presented as a sword in VR) and cut the fruits into pieces. The participants achieved higher scores with more fruit cut. Each trial lasted approximately 1.5 minutes.

**Koorring Wonderland—Mecadino’s Attack game package** Two games in this package were used. *Mirrorland’s Saber*: The participants needed to cut blocks flying to them into two pieces using the controller (presented as a red or blue sabre in VR) in the virtual environment; the colour of the blocks should match the colour of the sabre. *Mock Turtle’s Magic Shield*: The participants needed to block arrows flying toward them with the controller (presented as a red or green shield in VR) in the virtual environment; the colour of the arrows should match the colour of the shield. The participants won a gold, silver or bronze medal, depending on the number of blocks they cut or arrows they blocked successfully. Each trial lasted approximately 2.5 minutes.

**Koorring Wonderland—Heart Castle Crush game package** Three games in this package were used. *Humpty Dumpty*: The participants needed to hold the controller (presented as a red or blue hammer in VR) and whack the head of

the Humpty Dumpty running closer to them in the virtual environment. The Humpty Dumpty disappeared immediately after being whacked successfully. The colour of the Humpty Dumpty should match the colour of the hammer. *Garden to Live Flowers*: The participants stood in front of a tree in the virtual environment. The participants needed to hold the controller (presented as a paintbrush in VR) and paint the white flowers on the tree red. The painted flowers disappear immediately, and white flowers kept showing up. *Koorobo Toorobo*: In the virtual environment, there were four buttons of different colours (purple, green, yellow and red) in front of the participants, and different colours of blocks kept showing up in a circle in the distance. The participants needed hold the controller (presented as a virtual hand in VR) and press the corresponding button according to the colour of the block. The participants won a gold, silver or bronze medal, depending on the number of Humpty Dumpty they hit, flowers they painted or buttons they pressed successfully. Each trial lasted approximately 2.5 minutes.

**Space Slurpies** Food (e.g. the cube, tetrahedron and icosahedron) was floating in the air and moving around the participants in the virtual environment. The participants needed to hold and move the controller (presented as a snake in VR) to keep eating food, during which the snake grew in length and size. The participants achieved higher scores with more food ate. Each trial lasted approximately 2 minutes.

## Appendix 2. Distribution of the participants’ responses to the EQ-5D-5L.

EQ-5D-5L	Intervention			Control		
	Baseline (n = 25)	Week 1 (n = 23)	Week 2 (n = 14)	Baseline (n = 25)	Week 1 (n = 18)	Week 2 (n = 14)
<b>Mobility</b>						
No problems	2 (8.0%)	2 (8.7%)	1 (7.1%)	3 (12.0%)	4 (22.2%)	1 (7.1%)
Slight problems	8 (32.0%)	8 (34.8%)	5 (35.7%)	5 (20.0%)	6 (33.3%)	10 (71.4%)
Moderate problems	9 (36.0%)	12 (52.2%)	8 (57.1%)	10 (40.0%)	5 (27.8%)	2 (14.3%)
Severe problems	2 (8.0%)	1 (4.4%)	0 (0.0%)	5 (20.0%)	2 (11.1%)	1 (7.1%)
Unable to walk about	4 (16.0%)	0 (0.0%)	0 (0.0%)	2 (8.0%)	1 (5.6%)	0 (0.0%)
<b>Self-care</b>						
No problems	3 (12.0%)	1 (4.4%)	1 (7.1%)	1 (4.0%)	2 (11.1%)	1 (7.1%)
Slight problems	5 (20.0%)	7 (30.4%)	7 (50.0%)	6 (24.0%)	7 (38.9%)	7 (50.0%)

(continued)



