Effects of virtual reality in post-stroke aphasia: a systematic review and meta-analysis

Yun Cao[1,2] • Xing Huang[1,2] • Binlong Zhang[3] • Georg S. Kranz[4,5] • Danli Zhang[1,2] • Xiaolin Li[1,2] • Jingling Chang[2]

[1] Beijing University of Chinese Medicine, Beijing, China

[2] Department of Neurology, Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing, China

[3] Department of Acupuncture and Moxibustion, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing, China

[4] Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong, China

[5] The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong, China

Corresponding author: Prof. Jingling Chang, Department of Neurology, Dongzhimen Hospital, Beijing University of Chinese Medicine, No.5 Haiyuncang, Dongcheng District, Beijing 100700, China

E-mail: ear6979@163.com

Abstract

Objective To investigate whether virtual reality (VR) interventions have beneficial effects on the functional communication and language function of patients with post-stroke aphasia (PSA).

Methods We searched nine electronic literature databases and two clinical registry platforms to identify randomized controlled trials (RCTs) and quasi-RCTs performed up to September 2020. Screening, quality assessment, and data collection were performed by two authors independently, using standard protocols. Data aggregation and risk of bias evaluation were conducted using Review Manager Version 5.4. The quality of evidence was evaluated with GRADEpro.

Results A total of five studies involving 121 participants met the inclusion criteria and were appraised. Four studies were included in the quantitative synthesis. VR reduced the severity of language impairment with borderline significance [SMD (95%CI) = 0.68[-0.01, 1.36], P=0.05]. The meta-analysis showed no statistical difference in functional communication [SMD (95%CI) =0.41[-0.29, 1.12], P=0.25], word finding [SMD (95%CI) =0.42[-0.24, 1.08], P=0.21] and repetition [SMD (95%CI) =0.16[-0.62, 0.94], P=0.68] between VR group and the control group.

Conclusion This review demonstrated a borderline positive clinical effect of VR for the severity of language impairment when compared with conventional rehabilitation therapy. Conversely, VR had no effect on functional communication, word finding, and repetition. Further research is warranted to reach more definite conclusions.

Keywords Aphasia • Virtual Reality • Rehabilitation • Systematic review • Meta-analysis

Introduction

Stroke is a life-threatening disease with a high rate of disability and mortality [1-3]. Approximately one-third of people who suffer from a stroke develop aphasia [4-6], which significantly influences the individual, families, and communities. Identifying effective intervention is therefore vital. Speech and language therapy (SLT) is commonly used as the treatment methods for aphasia [7]. The latest Cochrane review on SLT of PSA concluded that therapy can improve functional language effectively [8]. Where therapy is intensive, high-repetition, dose-dependent, and task-oriented, patients can benefit from rehabilitation [9]. ***Several factors, however, reduced the chances of people with aphasia (PWA) receiving optimal services. The limited number of therapists, for example, cannot meet the demands of intensive treatment [10]. Furthermore, difficulties associated with transportation and high rehabilitation costs result in not all eligible patients being able to receive timely and effective treatments [11]. Therefore, it is critical to identify efficiency, convenient and home-based approaches for both enhancing and prolonging the benefits of aphasia intervention [12]. With the development of digital technology, VR has been explored as a mode of rehabilitation intervention and an alternative route for care delivery [13,14], which could add beneficial components to current rehabilitation***

strategies. Some studies have explored the applications of VR in PWA, highlighting feelings of fun and enjoyment [15], goal orientation [16], and multisensory stimulation at a high intensity [17]. These characteristics have a real potential to make VR technology an ideal tool in improving language function.

Recently, multiple clinical trials on the effects of VR training gave support to the use of VR in individuals PSA patients to improve functional communication and language function, especially in terms of linguistic level such as vocabulary. Language faculty are represented in a multimodal dimension in which word semantics contain sensorimotor properties, it depends on areas not previously hypothesized by the traditional approach, such as the sensorimotor network. One hypothesis was put forward that gestures participate in language production by increasing the semantic activation of words on the basis of sensory-motor features. Because the interactions of participants with VR scenes rely on the use of the handle [18], that is to say, gesture and language training will be carried out simultaneously during the operation. Therefore, it's easier to enhance sensorimotor properties in a VR environment compared to the usual speech and language therapy settings. Furthermore, it is reported that PWA have more difficulties in understanding verbs than nouns. These various studies have indicated that stronger neurological stimulation can provide the function of inducing verbal output. While the output of verbs seems to be better for dynamic stimuli than for static stimuli. In other words, the more dynamic the stimulus presentation, the closer it is to the real-life action word. In consequence, compared with the static image stimulus, the dynamic virtual reality stimuli that integrates vision, auditory and movements in training can help PWA improve verb production [19].

VR is now being used increasingly on patients with PSA [20]. ***There are many proponents of the view that VR offers a range of potential benefits for the treatment of aphasia, such as rich virtual***

environments, immediate feedback, and engaging experiences [21]. By offering a high-intensity task-oriented rehabilitation training, the interactivity and motivation of participants are facilitated [22]. In particular, when virtual therapy is provided via telehealth, it can enable patient access anytime and anywhere, assisting the therapist to keep track of the training situation with patients in real-time to adjust the rehabilitation program in time [23]. Furthermore, participants can get more feedback about the performance during VR training than in real-world practice [24]. During intervention, virtual therapy may help to reduce the feeling of embarrassment in real-world language function training. Relying on an immersive experience, VR may also greatly alleviate the shortage of therapists and continue the effectiveness of rehabilitation. The evidence for the use of VR in PSA has not, however, been systematically evaluated. It is essential to review the evidence of VR on the improvement of functional communication and language function in PWA to enable clinicians to have an up-to-date understanding of the effects of these clinical applications. In order to gain an increased understanding of the feasibility and limitations of VR in treating PSA and explore the direction of future study, we reviewed the literature and completed a meta-analysis to evaluate the effectiveness of VR applications for language training after stroke. ***The research question of this systematic review was formulated using PICOS (Population, Intervention, Comparison, Outcomes, and Study Design). The population was defined as people who had a diagnosis of aphasia following a stroke, which included different types of aphasia. The intervention was considered to be any types of virtual reality equipment and technologies useful for highlighting virtual engineering for the treatment of speech rehabilitation. The included studies were mainly concerned with the comparison of VR and the conventional speech therapy. For the outcomes, measures related to functional communication and language function. The study types included randomized controlled trials and quasi-randomized controlled trials.***

Methods

Search strategy

The protocol of this study was registered on PROSPERO International prospective register of systematic reviews (No. CRD42020169136). A search was made for RCTs and quasi-RCTs of VR for PSA, without language restriction. The following databases were searched from inception to September 2020: (1) MEDLINE; (2) EMBASE; (3) CENTRAL; (4) Web of Science; (5) Google Scholar; (6) SinoMed; (7) China National Knowledge Infrastructure (CNKI); (8) VIP Database for Chinese Technical Periodicals (VIP); (9) Wanfang Database. Ongoing and registered trials were searched on the following registers from inception to September 2020 to identify potential RCTs and quasi-RCTs: (1) ClinicalTrials.gov (http://clinicaltrials.gov/ct2/home); (2) ChiCTR (http://www.chictr.org.cn/index.aspx). Boolean search teams included the following: “aphasia,” “stroke,” “virtual reality,” “virtual reality exposure therapy,” and “computer simulation”. In PubMed, the searches were performed by using MeSH descriptors.

Selection criteria

The inclusion criteria for this review were as follows：(1) RCTs and quasi-RCTs that evaluated the clinical effect of VR on PSA with at least five participants. (2) The participants were diagnosed with stroke-induced aphasia. (3) There was a treatment group (isolated VR rehabilitation or in combination with other therapies) and a control group (rehabilitation therapy), other therapies should be same in the treatment and control groups. We included studies comparing VR conditions with controls (e.g., waitlist, placebo, treatment-as-usual). Articles were excluded if subject characteristics or interventions could not be accurately extracted.

Outcome measures

The primary objective was to evaluate the efficacy of VR compared with conventional intervention on functional communication, assessed on scales including the Amsterdam-Nijmegen Everyday Language Test (ANELT), Communicative Abilities of Daily Living (CADL), Communicative Activity Log (CAL), Communicative effectiveness Index (CETI). ***The important goal of linguistic rehabilitation is the improvement in functional communication which is the individual's ability to understand and convey information in everyday life situations. Therefore, such improvements are seen as the gold standard for demonstrating the effectiveness of interventions in PSA.***

The secondary outcomes were language functions assessed on scales including the Boston Diagnostic Aphasia Examination (BDAE) severity classification, Bilingual Aphasia Test (BAT), China Rehabilitation Research Center Aphasia Examination (CRRCAE), Profile of Word Errors and Retrieval in speech (Powers), and Aphasia Battery of Chinese (ABC), ***Aachen Aphasia Test (A.A.T.)***. Language function involves the aspects of spontaneous speech, severity of language impairment, auditory comprehension, repetition, naming, reading and writing.

Data extraction

Data from each study were independently extracted by two reviewers (YC and XH) using a standard data recording which included the study design, number of subjects, mean age, stroke duration, treatment protocol, dropout number, outcome measures, pre- and post-treatment means and standard deviations for outcome measures. The authors were also contacted via email to get missing data if possible.

Quality assessment

The risk of bias was assessed using the Cochrane Collaboration "Risk of bias" assessment tool by the reviewers [25]. Any disagreement was resolved by consensus. According to the recommendations of the Cochrane Handbook, we also applied the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach to evaluate the strength of the main outcomes. GRADEpro version 3.6.1 software was used for data analysis and synthesis.

Statistical analysis

All the data in this analysis were continuous variables. A mean difference (MD) with 95% confidence interval (CI) was used to calculate treatment effects of outcomes measured by the same scale. A standardized mean difference (SMD) with 95% CI was used to pool the data because different scales were used to evaluate linguistic outcomes. The heterogeneity across each effect size was evaluated with Q-statistics and the I^2 index, and an I^2 value of more than 50% indicated statistical heterogeneity. The fixed-effect model (FEM) was chosen to synthesize data without significant heterogeneity and a random-effects model (REM) was used for data with significant heterogeneity in the meta-analysis. Statistical calculations were performed using Review Manager 5.4 (The Cochrane Collaboration, Copenhagen, Denmark, 2014). No subgroup analysis was performed according to participant characteristics (e.g., duration of stroke) or type of VR technology (immersive VR versus non-immersive VR) due to the low number of RCTs included in the study.

Results

Study selection

The search strategy identified 896 records from the database. After the exclusion of ineligible studies, five studies were selected and reviewed for the qualitative synthesis of evidence. However, due to the high heterogeneity between one research and the other four studies in outcome indicators, the effect sizes were not combined quantitively in a meta-analysis. The flow of reference is shown in **Figure 1**.

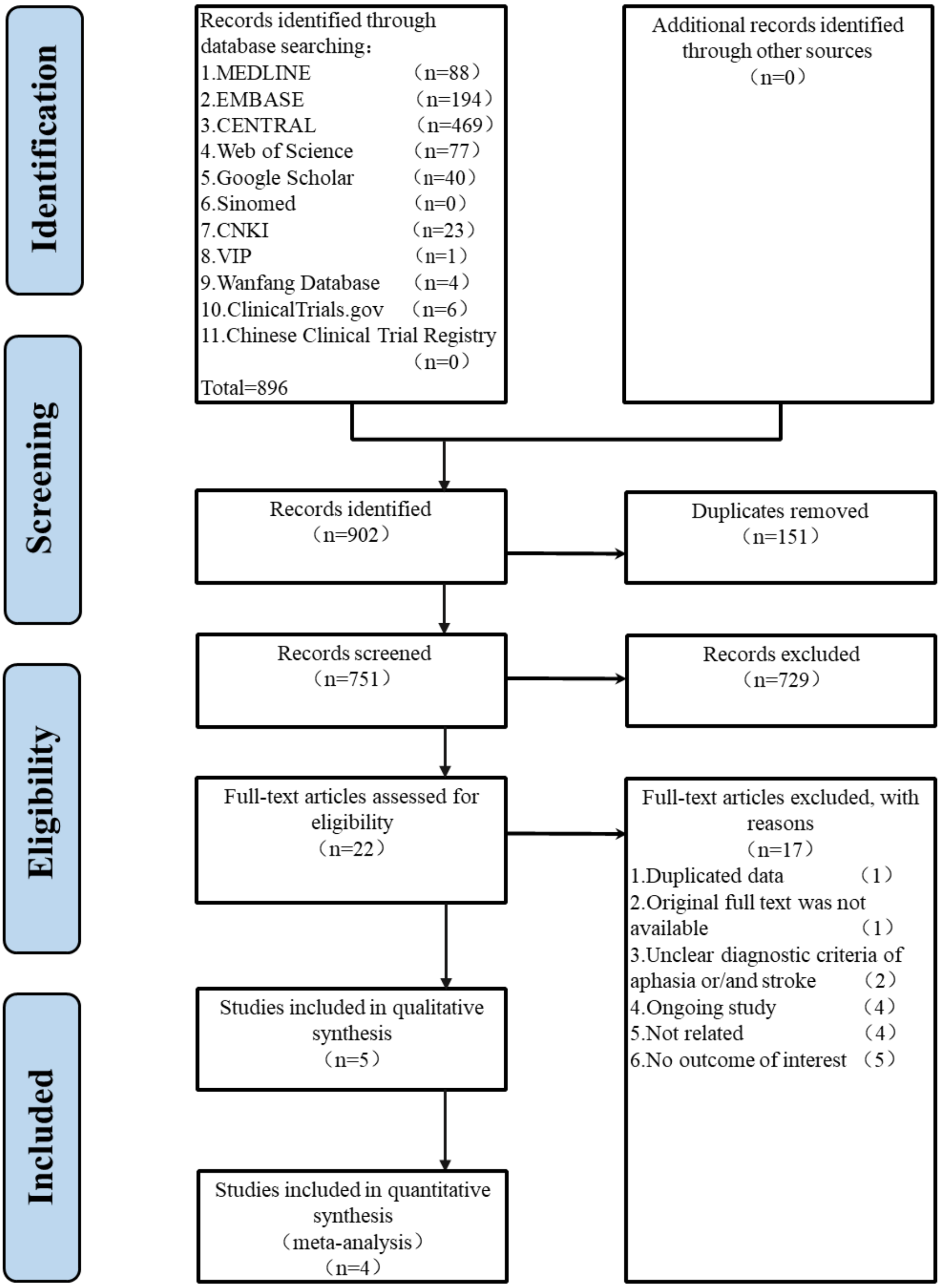

Figure 1. Study flow diagram

Study design and characteristics

Each of the five studies compared a different type of VR technology to a control group. The training contents of the VR equipment were the same as conventional speech function training in the majority of the studies, which included an exercise on speech/language abilities, memory, attention, etc. VR intervention across the entire study ranged from 10 hours to 96 hours, and the duration of each session ranged from 30 minutes to 60minutes. Intervention length ranged from 4 weeks to 24 weeks, and treatment frequency ranged from twice a week to five days a week. Four trials employed randomized parallel-group methods, only one followed a cross-over design. Detailed information on the characteristics of included studies is shown in **Table 1**.

Table 1. The characteristics of included studies										
Study	**Number of participants**	**Participants**		**Age (Years)**		**Time Post Onset (Month)**		**Intervention**	**Therapy duration**	**Outcome measures**
		Type of Stroke	**Type of Aphasia**	**Treatment**	**Control**	**Treatment**	**Control**			
Marshall et al. [16] 2016	20	Unclear	Unclear	59.00±13.61	56.6±9.73	70.10±68.91	54.1±34.46	1 The EG received EVA Park intervention in week 8-12, and no further intervention in weeks 8-12. 2 The CG received EVA Park intervention in weeks 2-6, and no further intervention in weeks 7-13.	5 weeks	1 CADL-2 2 Verbal Fluency 3 Conversation% content words 4 Conversation content words/turn 5 Narrative words per minute 6 Narrative sentences# 7 CCRSA 8 Friendship Scale
Grechuta et al. [26] 2019	17	1 ischemic stroke 2 hemorrhagic stroke	Broca's Aphasia	55.66± 8.40	53.5±11.33	61.66±46.89	58±52.04	1 The EG received the training of RGSa. 2 The CG received standard SLT targeting specific linguistic deficits in a therapist-patient setting.	8 weeks	1 BDAE 2 CAL 3 VocabT 4 FMA-UE
Maresca et al. [27] 2019	30	1 ischemic stroke 2 hemorrhagic stroke	Unclear	51.1±10.3	51.4±12.7	Unclear	Unclear	1 The EG underwent an ELT performed using the VRRS(T0-T1) and were provided with touchscreen tablet (VRRS-Tablet) in the second phase(T1-T2). 2 The CG were trained with a traditional linguistic treatment(T0-T1). In the second phase (T1-T2), the	24 weeks	1 TT 2 ADRS 3 EQ-5D 4 ENPA

								CG was delivered to territorial services, where they undergo conventional speech therapy.		
Giachero et al. [28] 2020	36	1 ischemic stroke 2 hemorrhagic stroke	Nonfluent aphasia	Unclear	Unclear	36.33±9.86	49.17±10.57	1 The EG underwent conversational therapy during VR everyday life setting observation. 2 The CG was trained in a conventional setting without VR support.	24 weeks	1 A.A.T. 2 C.A.P.P.A. test 3 VASES 4 WHOQoL Questionnaire
Zhang et al. [29] 2017	18	1 ischemic stroke 2 hemorrhagic stroke	Broca's Aphasia	45.11±15.34	42.33±15.86	13.11±6.45	14.56±5.41	1 The EG was treated with conventional speech function training for 20 minutes and same training content with control group. After 20 minutes, the experimental group was required to wear VR equipment. 2 The CG was treated with conventional speech function training for 40 minutes.	4 weeks	1 CRRCAE 2 BDAE

EG [Experimental Group]; CG [Control Group]; SLT [Speech and language therapy]; ELT [experimental linguistic treatment]; VRRS [VR rehabilitation system]; RGSa [Rehabilitation Gaming System for Aphasia].

Quality assessment

Figure 2 provided an overview of the methodological quality of the included papers. For the generation of random sequences, three studies were evaluated as low RoB because the specific methods were used, including random number table, computer-generated stratified randomization, and Research Randomizer. Unclear allocation concealment was the main cause of selection bias. For incomplete outcome data, all the studies did not provide enough information to assess this domain due to the lack of mention of Intention-To-Treat (ITT) analysis or Pre-Protocol (PP) analysis. For blinding, because of the nature of the VR interventions, most of the studies did not report blinding the participants. All of the included studies were evaluated at low risk of reporting bias. Three studies were regarded free from other sources of bias due to the balance baseline characteristics (age, sex, education).

The quality of the evidence was graded as 'low' for the outcomes of functional communication, word finding, the severity of language, and repetition. According to the risk of bias summary and graph (**Figure 2**), the risk of bias in several studies was unclear due to the poor reporting and selection ambiguity. The limitations of study design and publication bias further contributed to the evidence which was rated as 'low' quality when rated using the GRADE system. The GRADE evidence profiles are shown in **Table 2**.

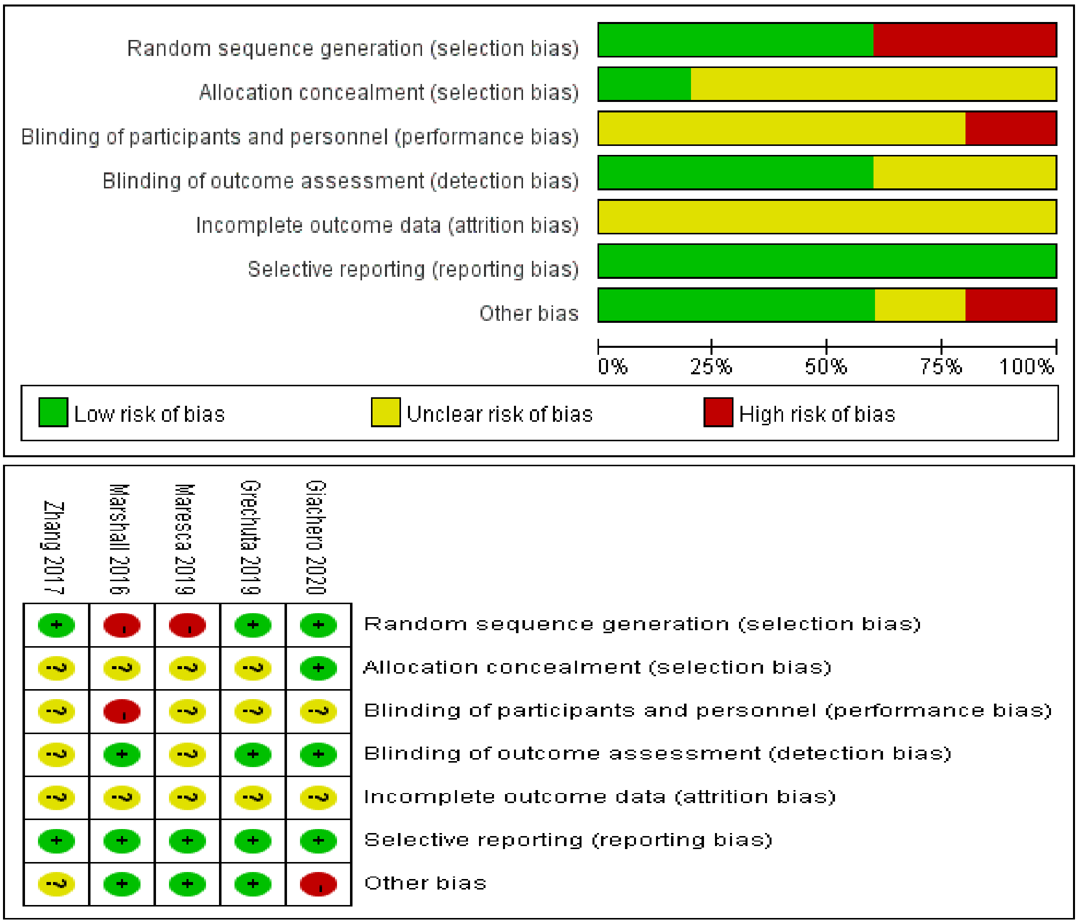

Figure 2. Risk of bias graph and summary. "+": low risk; "?": unclear risk; "–": high risk. The source of risk of bias consists of selection, detection, reporting, and other bias

Table 2. GRADE evidence profile of VR VS Language rehabilitation therapy												
Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Virtual Reality	Speech and Language Rehabilitation	Relative (95% CI)	Absolute		
Functional communication (measured with: CADL, CAL; Better indicated by lower values)												
2	randomized trials	serious[1,2,3]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[4]	19	18	-	SMD 0.42 higher (0.24 lower to 1.08 higher)	⊕⊕OO **LOW**	CRITICAL
Word finding (measured with: VocabT, POWERS; Better indicated by lower values)												
2	randomized trials	serious[1,2,3]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[4]	19	18	-	SMD 0.42 higher (0.24 lower to 1.08 higher)	⊕⊕OO **LOW**	CRITICAL
Severity of language (measured with: BDAE; Better indicated by lower values)												
2	randomized trials	serious[2,3]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[4]	18	17	-	SMD 0.68 higher (0.01 lower to 1.36 higher)	⊕⊕OO **LOW**	CRITICAL
Repetition (measured with: AAT, CRRCAE; Better indicated by lower values)												
2	randomized trials	serious[2,3]	no serious inconsistency	no serious indirectness	no serious imprecision	reporting bias[4]	27	27	-	SMD 0.09 higher (0.45 lower to 0.63 higher)	⊕⊕OO **LOW**	CRITICAL

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

[1] No details of random protocol were reported; [2] Lack of allocation concealment; [3] Didn't report the implementation of blinding; [4] Quantitative evaluation of the included data indicated publication bias

Main results

Efficacy of VR regarding functional communication

The primary outcome was functional communication, which was measured by CADL and CAL in two studies. A total of 47 participants were enrolled in this comparison [16,26]. As displayed in **Figure 3-a**, there existed no significant differences between VR and the control groups (P=0.25; SMD 0.41, 95%CI -0.29 to 1.12). The meta-analysis revealed low statistical heterogeneity (I^2=13%; P=0.28).

Efficacy of VR regarding language functions

Secondary outcomes were the severity of language impairment (measured by BDAE severity classification) and language abilities, including word finding, spontaneous speech, auditory comprehension, repetition, naming, reading, and writing. Due to the differences in evaluation methods, however, with the exception of word finding, repetition and severity of language, it was not appropriate to pool the studies for other outcome measures. There was no significant difference of word finding between VR and the control group by using a REM (P=0.21; SMD 0.42, 95%CI -0.24 to 1.08) (**Figure 3-b**). No statistical heterogeneity was found among included studies (I^2=0%; P=0.42). As shown in **Figure 3-c**, no significant difference was found in the repetition comparison (P=0.68; SMD 0.16, 95%CI-0.62 to 0.94). We also found medium heterogeneity among studies as the value of I^2 was 47%. As shown in the **Figure 3-d**, analysis of data from 35 participants showed the borderline significance between VR and the control group on the outcome of the severity of language (P=0.05; SMD 0.68, 95%CI-0.01 to 1.36). No statistical heterogeneity was detected among included studies (I^2=0%; P=0.96).

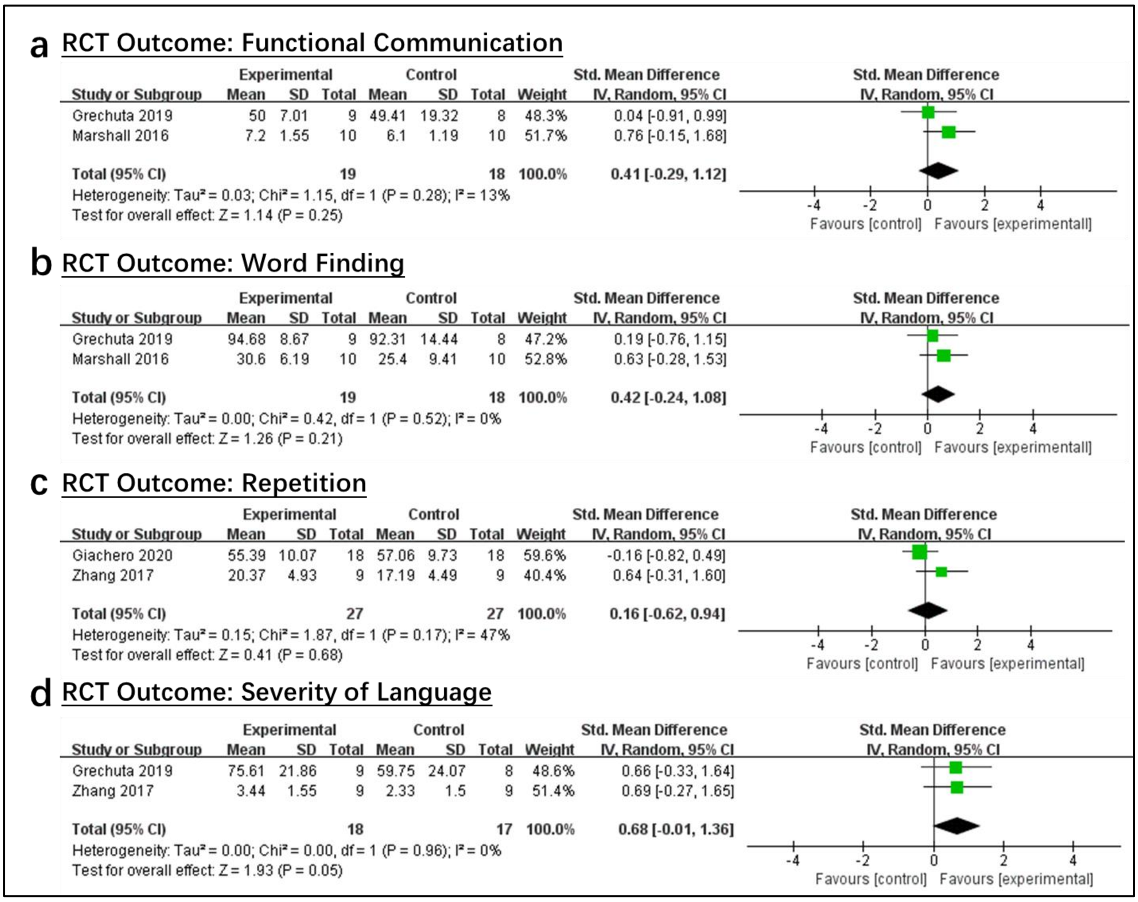

Figure 3. Meta-analysis of RCTs using VR in Functional Communication (a)，Word Finding (b)，Repetition (c), Severity of Language (d).

Discussion

To our knowledge, this is the first meta-analysis to evaluate the effects of VR compared with SLT on PSA. A primary goal of aphasia intervention is to improve daily communication, hence functional communication is considered as our primary outcome. However, no evidence of differences is found from the primary language outcomes in the meta-analysis. The stroke guidelines acknowledged the benefits of SLT for people with aphasia following stroke. Functional communication was significantly improved in PWA who received SLT at high intensity [7,30]. As an emerging and novel therapeutic approach, VR for aphasia is still at an early stage. There is still room for VR improvement when compared with conventional effective therapies. Although VR did not show significant advantages for functional communication, the results would be accepted. For the severity of language, a borderline significance was found. The relatively consistent results showed a significant twofold relationship between language impairment and functional communication measures [31]. Furthermore, the severity of language impairment seemed to be one of the best predictors of aphasia outcome [32]. Therefore, the severity of language impairment is taken as the second outcome. An extensive body of research argued that the benefits of high-intensity approaches to SLT were related to the severity of impairment [33]. Although the result was borderline meaning, it seemed to reflect on the potential of VR for the severity of language impairment. When there is a clinically meaningful treatment effect, borderline P values usually appear due to an insufficient number of participants or events [34]. Moreover, borderline P values can also occur when the treatment effect is less than expected, which required a large trial to produce a P value<0.05 [35]. Therefore, it is necessary to carry out well-designed research and conduct trials with a larger sample before drawing conclusions.

For methodological quality, RoB tools were used to facilitate the improved appraisal of evidence. The overall methodological quality of the included studies was predominantly poor, largely due to the high RoB. For the reason of the small number of studies, it is not possible to rule out or confirm the presence of possible publication bias, as an assessment of funnel plots were feasible in reviews with ten or more studies. Moreover, the quality of evidence was graded by implementing the GRADEpro in an attempt to verify the strength of the recommendation. In this study, the quality of evidence was considered to be "low" for all outcomes (Weak recommendation). In the absence of complete and definitive evidence, consistent recommendations can be reached in the future through the strict expert consensus method [36]. The Consensus Development Conference (CDC), Nominal Group Technique (NGT), and Delphi method are the three specific expert consensus methods that can be used commonly in the medical field [37]. Future trails need robust study design with a large sample size, and may also consider registering and publishing protocols to promote proper research implementation and reporting [38].

Clinical evidence suggested that the therapy of PSA should satisfy the need of high-intensity, socially embedding, and goal-oriented [39]. Based on the above principles, VR intervention can exactly meet these requirements. It can provide interesting, immersive experience which can enhance the learning motivation and encourage patients to promote intensive language practice. VR provides individualized training with appropriate intensity and difficulty according to the characteristics and the needs of the patients [40]. It is well known that extrinsic feedback [41], especially visual feedback, which can improve the learning rate [42] and facilitate corticospinal activation to a greater extent [43], so patients with stroke can benefit from practice with augmented feedback [44]. The multisensory stimulation of VR training promotes the recovery of functional communication, mnestic-attentive functions, visuospatial cognition, and behavior abilities of patients with neurological disorders [45]. Neuroplasticity and functional reorganization are the main mechanisms of conventional clinical practice for stroke rehabilitation, which are utilized by the brain to encode experiences and learn new behaviors [46]. VR can provide a new strategy to improve and amplify the neural plasticity process [47], which may be related to the reactivation of brain neurotransmitters due to the training under the virtual environment [48].

In this review, we treated VR as a homogenous intervention, although the environments and hardware of VR used differently. Due to the complexity and variety of the study design, population, and procedure type, it is unlikely for us to separate differences between the various VR types because of insufficient statistical power. The available evidence was limited, and the quality of evidence included in studies was moderate to low. Although the application of VR in post-stroke rehabilitation

has potential, some limitations need to be acknowledged.

Participants. In view of the small sample size of individual test, the inclusion of diverse stroke population was not strictly limited. We did not restrict the following strictly: the stroke type (ischemic or hemorrhage), aphasia type (fluent aphasia, non-fluent aphasia) [49], stroke locality (cortical or subcortical), level of aphasia severities (mild, moderate, or severe), VR type (immersive, non-immersive, or semi-immersive) and stage of recovery.

Interventions. VR ranges from non-immersive to fully immersive, depending on the degree to which the users are isolated from the physical surroundings when interacting with the virtual environment. Semi-immersion and non-immersion techniques have been widely used because most studies required the assistance of therapists. People under the test could perform activities and explore both in the real and virtual environment, but the feeling of "being there" would be reduced. The types of VR systems also varied greatly, ranging from commercial game systems to engineer-built systems. In order to obtain more complete data, the inclusion criteria did not limit such differences, which made comparisons between these interventions more difficult. Although the frequency and intensity of the intervention were described in various studies, the specific contents of the intervention were not explained in detail, which made it difficult to repeat the included studies. Marshall et al. [16] designed a semi-immersive virtual world called EVA Park, which was developed through a process of codesign with five aphasia patients. It was a virtual island with many facilities such as houses, cafés, restaurants, health centers, hair salons, tropical bars and discos. The environment was interactive. Users could use headsets and microphones to communicate with other users with personalized avatars in real time through voice. They could also navigate their avatars via a key pad and a mouse. The content of intervention was mainly determined by participants who set at least three goals. Grechuta et al. [26]

investigated a new VR intervention which was called the Rehabilitation Gaming System for aphasia (RGSa), it could provide lexical and syntactic training in a multimodal, goal-oriented manner in the context of dyadic peer-interaction. Two patients interacted with the system by performing planar arm movements which were tracked and mapped onto the upper limbs of the avatar, thus providing embodiment and allowing the interaction with virtual objects. The paradigm requires participation in daily communication acts. The RGSa sessions were supervised by a therapy assistant who did not offer any other services. Maresca et al. [27] investigated a touchscreen tablet (VRRS-Tablet) which contained about 30 different exercises. These exercises were specifically as follows, attention, memory, perception, executive functions, and speech/language skills. VRRS had two main categories of exercises. The first category included 2D exercises, where patients could interact with objects and scenarios through the touch screen or a particular magnetic sensor coupled with buttons. The second one consisted of 3D exercises, where patients could interact with 3D virtual scenarios and immersive objects through a magnetic localization sensor generally positioned on the hands. The program was designed to restore the language skills by naming characters displayed on the screen, as well as tasks of composition, writing and rewriting suggested by acoustic, textual or visual item. Giachero et al. [28] made full use of semi-immersive VR environment to investigate its therapeutic benefits, mainly in enhancing language skills, communication efficacy, and psychosocial treatment aspects. The semi-immersive VR scenarios were projected through a screen and represent daily communication situations through different cognitive exercises which referred to language, memory, and attentional tasks. Meanwhile, the interaction between patients was carried out by a speech therapist. The VR equipment used by Zhang et al. [29] was a mirror neuron rehabilitation training system, which consisted of control unit, panoramic helmet and mirror neuron rehabilitation training software, etc.

The training contents of VR was as follows: noun listening and reading, verb listening and reading, phrase listening and reading, sentence listening and reading.

Comparison. Four of the included studies were of a parallel design and had a control group. The control group treated with conventional speech function training, including naming, repetition, spelling and so on. The intensity and frequency of the training in both EG and CG were identical.

Outcomes. The outcome measurements used across the studies were inconsistent. None of the five studies used all the outcomes, which highlighted a gap between the available evidence and the literature on the subject. Specific studies focused on single rather than multiple dimensions. Because of the different languages in the research and the cultural background, the linguistic scales selected by various researches might also be different, which is mainly due to the structural features of language, clinical experience and different perspectives of language phenomena. The formation of linguistic scales is usually drawn from psycholinguistics, neurolinguistics, cognitive neuropsychology, and further involves the combination of listening, speaking, reading, writing, calculation, and other aspects. There may not be a clinical linguistic scale that can simultaneously assess multiple languages at present. Therefore, it is of clinical significance to explore sensitivity and specificity of the linguistic scales for the diagnosis and assessment of PSA.

Study Design. The included studies employed randomized controlled designs except for one quasi-randomized study. The randomization methods were described in detail in three studies, which could avoid selectivity bias to some extent. It is impossible to blind participants to the treatment allocation due to the characteristics of VR training, which might introduce bias into the estimation of treatment effects. The outcome assessors were blinded in most studies.

Limitations

The current evidence is the first step in evaluating the effects of VR therapy. ***It was possible that relevant studies were missed despite a comprehensive search strategy across multiple databases.*** During the literature screening process, we found that many studies had been implemented without a rigorous experimental design. The majority of the studies were carried out as case studies and semi-structured interviews. ***The meta-analysis should be different according to the clinical syndrome, and we have recognized this as an important limitation. Due to the lack of lack of standardized clinical studies with large sample size, it is difficult to carry out this comparative analysis comprehensively.***

Future trials

The GRADE criteria proved that the overall evidence grade of outcomes was low. Thus, it is suggested that future trials should use well-designed RCTs in accordance with the recommendations of the Consolidated Standards of Reporting Trials (CONSORT) [50]. On the basis of previous observations, the following proposals and suggestions are provided for future clinical trials: (1) In order to minimize or avoid performance bias caused by open-label trials, the use of subjective evaluation indicators should be avoided as much as possible. (2) Future investigations can consider different types, severities, and evolution of PSA in order to expand the scope of application in the field of VR. (3) Further researches can focus on the effectiveness of different VR modes, and search for the best VR mode which is also suitable for the treatment of PSA. (4) A control group should be arranged as well and given conventional speech training which keeps the same intensity, frequency, and training materials as the VR group. (5) Researchers could study the practicability and confirm the adverse reaction of VR by collecting follow-up data and using well-validated measures of language disorder outcomes.

Conclusion

According to the current available RCTs using VR with PWA, VR shows promise as an effective approach in treating PSA through improving functional communication and language function. Notwithstanding the present results are uncertain about the effectiveness of VR for PSA, they may provide some direction for facilitating future research and clinical practice in this field. Because of the limitation of current evidence, high-quality RCTs with improved allocation concealment and blind design are needed to promote the knowledge about the optimal rehabilitation strategy for PWA.

Acknowledgements We would like to thank Professor Anne Whitworth for her great help in improving this manuscript.

Authors' contributions YC and XH identified and screened the articles, summarized and analyzed the data. The discrepancies were resolved through discussion with BZ. YC wrote the initial version of manuscript. All authors read, edited and approved the final manuscript.

Funding information This work was supported by the National Natural Science Foundation of China (grant number 81973790) and the Beijing Municipal Natural Science Foundation (grant number 7181005).

Availability of data and material No additional data are available.

Compliance with ethical standards

Ethical approval and consent to participate Not applicable.

Consent for publication All authors have read and approved the manuscript for publication.

Conflict of Interests Statements The authors declare that they have no conflict of interest.

Statement of human and animal rights This article does not contain any studies with animals performed by any of the authors.

ORCID iD

https://orcid.org/0000-0001-7716-2997

https://orcid.org/0000-0002-7881-8174

References

1. Murray CJL, Barber RM, Foreman KJ, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: quantifying the epidemiological transition[J]. The Lancet, 2015, 386: 2145-2191. https://doi.org/10.1016/s0140-6736(15)61340-x
2. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010[J]. The Lancet, 2012, 380: 2095-2128. https://doi.org/10.1016/s0140-6736(12)61728-0
3. Murray CJL, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010[J]. The Lancet, 2012, 380: 2197-2223. https://doi.org/10.1016/s0140-6736(12)61689-4 \
4. Berthier ML. Poststroke aphasia: epidemiology, pathophysiology and treatment[J]. Drugs & aging, 2005, 22(2): 163-182. https://doi.org/10.2165/00002512-200522020-00006
5. Engelter ST, Gostynski M, Papa S, et al. Epidemiology of Aphasia Attributable to First Ischemic Stroke: Incidence, Severity, Fluency, Etiology, and Thrombolysis[J]. Stroke, 2006, 37(6): 1379-1384.doi: https://doi.org/10.1161/01.str.0000221815.64093.8c
6. Béjot Y, Rouaud O, Jacquin A, et al. Stroke in the very old: incidence, risk factors, clinical features, outcomes and access to resources–a 22-year population-based study[J]. Cerebrovascular Diseases, 2010, 29(2): 111-121. https://doi.org/10.1159/000262306
7. Bakheit AMO, Shaw S, Barrett L, et al. A prospective, randomized, parallel group, controlled study of the effect of intensity of speech and language therapy on early recovery from poststroke aphasia[J]. Clinical Rehabilitation, 2007, 21(10): 885-894.

https://doi.org/10.1177/0269215507078486

8. Brady MC, Kelly H, Godwin J, et al. Speech and language therapy for aphasia following stroke[J]. Cochrane database of systematic reviews, 2012;5:CD000425. https://doi.org/10.1002/14651858.cd000425

9. Weiss PL, Keshner EA, Levin MF. Virtual Reality for Physical and Motor Rehabilitation[J]. Virtual Reality Technologies for Health & Clinical Applications, 2014. https://doi.org/10.1007/978-1-4939-0968-1.

10. Gerber SM, Schütz N, Uslu AS, et al. Therapist-Guided Tablet-Based Telerehabilitation for Patients With Aphasia: Proof-of-Concept and Usability Study[J]. JMIR rehabilitation and assistive technologies, 2019, 6(1): e13163. https://doi.org/10.2196/13163

11. Blonski DC, Covert M, Gauthier R, et al. Barriers to and facilitators of access and participation in community-based exercise programmes from the perspective of adults with post-stroke aphasia[J]. Physiotherapy Canada, 2014, 66(4): 367-375. https://doi.org/10.3138/ptc.2013-70

12. Sheehy L, Taillon-Hobson A, Sveistrup H, et al. Home-based virtual reality training after discharge from hospital-based stroke rehabilitation: a parallel randomized feasibility trial[J]. Trials, 2019, 20(1): 333. https://doi.org/10.1186/s13063-019-3438-9

13. ***Triegaardt, J., Han, T.S., Sada, C. et al. The role of virtual reality on outcomes in rehabilitation of Parkinson's disease: meta-analysis and systematic review in 1031 participants. Neurol Sci 41, 529–536 (2020). https://doi.org/10.1007/s10072-019-04144-3***

14. Simone LK, Schultheis MT, Rebimbas J, et al. Head-mounted displays for clinical virtual reality applications: pitfalls in understanding user behavior while using technology[J]. CyberPsychology & Behavior, 2006, 9(5): 591-602. https://doi.org/10.1089/cpb.2006.9.591

15. Amaya A, Woolf C, Devane N, et al. Receiving aphasia intervention in a virtual environment: the participants' perspective[J]. Aphasiology, 2018, 32(5):538-558. https://doi.org/10.1080/02687038.2018.1431831
16. Marshall J, Booth T, Devane N, et al. Evaluating the Benefits of Aphasia Intervention Delivered in Virtual Reality: Results of a Quasi-Randomised Study[J]. PLoS ONE, 2016, 11(8): e0160381. https://doi.org/10.1371/journal.pone.0160381
17. Törnbom K, Danielsson A. Experiences of treadmill walking with non-immersive virtual reality after stroke or acquired brain injury–A qualitative study[J]. PloS one, 2018, 13(12): e0209214. https://10.1371/journal.pone.0209214
18. Li X , Zhang J , Cheng J , et al. Simulation of Switching Operation Based on Virtual Reality Method[J]. Journal of Physics: Conference Series, 2020, 1650(3):032121 (6pp). https://10.1088/1742-6596/1650/3/032121
19. Park H J , Oh C , Kwon S B . The Effect of Augmented Reality Contents on Verb Naming in People With Aphasia[J]. Journal of speech-language & hearing disorders, 2020, 29(3):27-34. https://10.15724/jslhd.2020.29.3.027
20. Antkowiak D, Kohlschein C, Kroo R, et al. Language therapy of aphasia supported by augmented reality applications[C]// 18th International Conference on E-health Networking, Application and Services (IEEE Healthcom '16). IEEE, 2016:1-6. https://doi.org/10.1109/healthcom.2016.7749511
21. Burdea G, Coiffet P. Virtual Reality Technology[J]. Presence, 2003, 12(6): 663-664. https://doi.org/10.1162/105474603322955950
22. Wang Z, Wang P, Xing L, et al. Leap Motion-based virtual reality training for improving motor functional recovery of upper limbs and neural reorganization in subacute stroke patients[J]. Neural

regeneration research, 2017, 12(11): 1823-1831. https://doi.org/10.4103/1673-5374.219043

23. de Rooij, I. J. M., van de Port, I. G. L., and Meijer, J. W. G. (2016). Effect of virtual reality training on balance and gait ability in patients with stroke: systematic review and meta-analysis. Phys. Ther. 96(12):1905–1918. https://doi.org/10.2522/ptj.20160054
24. Simonsen D, Popovic MB, Spaich EG, et al. Design and test of a Microsoft Kinect-based system for delivering adaptive visual feedback to stroke patients during training of upper limb movement[J]. Medical & biological engineering & computing, 2017, 55(11): 1927-1935. https://10.1007/s11517-017-1640-z
25. Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
26. Grechuta K, Rubio Ballester B, Espín Munne R, et al. Augmented dyadic therapy boosts recovery of language function in patients with nonfluent aphasia: a randomized controlled trial[J]. Stroke, 2019, 50(5): 1270-1274. https://doi.org/10.1161/STROKEAHA.118.023729
27. Maresca G, Maggio MG, Latella D, et al. Toward Improving Poststroke Aphasia: A Pilot Study on the Growing Use of Telerehabilitation for the Continuity of Care[J]. Journal of Stroke and Cerebrovascular Diseases, 2019, 28(10): 104303. https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.104303
28. Giachero, A., Calati, M., Pia L., et al., et al. Conversational Therapy through Semi-Immersive Virtual Reality Environments for Language Recovery and Psychological Well-Being in Post Stroke Aphasia. Behavioural Neurology, 2020, 1–15. https://doi.org/10.1155/2020/2846046
29. Zhang Y, Chen P, Li X, et al. Clinical research on therapeutic effect of virtual reality technology on

Broca Aphasia patients[J]. 2017:1-5. https://doi.org/10.1109/INCIT.2017.8257880

30. Bhogal S, Teasell R, Speechley M. Intensity of aphasia therapy, impact on recovery. Stroke. 2003; 34 (4): 987–993. https://doi.org/10.1161/01.STR.0000062343. 64383.D0

31. Irwin WH, Wertz RT, Avent JR. Relationships among language impairment, functional communication, and pragmatic performance in aphasia[J]. Aphasiology, 2002, 16(8): 823-835. https://doi.org/10.1080/02687030244000086

32. Osa García A, Brambati S M, Brisebois A, et al. Predicting Early Post-stroke Aphasia Outcome From Initial Aphasia Severity[J]. Frontiers in neurology, 2020, 11: 120. https://doi.org/10.3389/fneur.2020.00120

33. Stahl B, Mohr B, Dreyer FR, et al. Using language for social interaction: communication mechanisms promote recovery from chronic non-fluent aphasia[J]. Cortex, 2016, 85: 90-99. https://doi.org/10.1016/j.cortex.2016.09.021

34. Moher D, Liberati A, Tetzlaff J, et al. Research methods and reporting[J]. Bmj, 2009, 8: 332-336. https://doi.org/10.1136/bmj. b2535

35. Hackshaw A, Kirkwood A. Interpreting and reporting clinical trials with results of borderline significance[J]. BMJ, 2011, 343: d3340-d3340. https://doi.org/10.1136/bmj. d3340

36. Powell C. The Delphi techniques: Myth and realities. J Adv Nurs 2003; 41(4):376-382. https://doi.org/10.1046/j.1365-2648.2003. 02537.x

37. Mcmillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques[J]. International Journal of Clinical Pharmacy, 2016, 38(3):655-662. https://doi.org/10.1007/s11096-016-0257-x

38. Chan A W, Tetzlaff J M, Gøtzsche P C, et al. SPIRIT 2013 explanation and elaboration: guidance

for protocols of clinical trials[J]. Bmj, 2013, 346: e7586. https://doi.org/10.1136/bmj. e7586

39. Hilari K, Needle JJ, Harrison L. What are the important factors in health-related quality of life for people with aphasia? A systematic review[J]. Archives of physical medicine and rehabilitation, 2012, 93(1): S86-95. https://doi.org/10.1016/j.apmr.2011.05.028

40. Merians A S, Poizner H, Boian R, et al. Sensorimotor training in a virtual reality environment: does it improve functional recovery poststroke? [J]. Neurorehabilitation and neural repair, 2006, 20(2): 252-267. https://doi.org/10.1177/1545968306286914

41. Walker ER, Hyngstrom AS, Schmit BD. Influence of visual feedback on dynamic balance control in chronic stroke survivors. J Biomech, 2016;49(5):698-703. https://doi.org/10.1016/j.jbiomech.2016.01.028

42. Kang YJ, Park HK, Kim HJ, et al. Upper extremity rehabilitation of stroke: facilitation of corticospinal excitability using virtual mirror paradigm[J]. Journal of neuroengineering and rehabilitation, 2012, 9(1): 71. https://doi.org/10.1186/1743-0003-9-71

43. Valdés BA, Schneider AN, Van der Loos HFM. Reducing trunk compensation in stroke survivors: a randomized crossover trial comparing visual and force feedback modalities[J]. Archives of physical medicine and rehabilitation, 2017, 98(10): 1932-1940. https://doi.org/10.1016/j.apmr.2017.03.034

44. Riva G, Wiederhold BK, Mantovani F. Neuroscience of virtual reality: From virtual exposure to embodied medicine[J]. Cyberpsychology, Behavior, and Social Networking, 2019, 22(1): 82-96. https://doi.org/10.1089/cyber.2017.29099.gri

45. Maggio MG, Maresca G, De Luca R, et al. The growing use of virtual reality in cognitive rehabilitation: fact, fake or vision? A scoping review[J]. Journal of the National Medical Association,

2019, 111(4): 457-463. https://doi.org/10.1016/j.jnma.2019.01.003

46. Kleim JA, Jones TA. Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. J Speech Lang Hear Res. 2008;51(1): S225-239. https://doi.org/10.1044/1092-4388(2008/018)
47. Veerbeek JM, van Wegen E, van Peppen R, et al. What is the evidence for physical therapy poststroke? A systematic review and meta-analysis. PLoS One. 2014;9(2): e87987. https://doi.org/10.1371/journal.pone.0087987
48. Huet M, Jacobs DM, Camachon C, et al. Self-Controlled Concurrent Feedback Facilitates the Learning of the Final Approach Phase in a Fixed-Base Flight Simulator[J]. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2009, 51(6):858-871. https://doi.org/10.1177/0018720809357343
49. Ferro J M, Madureira S. Aphasia type, age and cerebral infarct localisation[J]. Journal of neurology, 1997, 244(8): 505-509. https://doi.org/10.1007/s11517-017-1640-z
50. Schulz, K.F., Altman, D.G. & Moher, D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. Trials 11,32(2010). https://doi.org/10.1186/1745-6215-11-32