

## RESEARCH ARTICLE

# Comparing two identically protocolized, multicentre, randomized controlled trials on caregiver-mediated exercises poststroke: Any differences across countries?

Marijn Mulder<sup>1,2</sup>, Rinske H. M. Nijland<sup>2\*</sup>, Judith D. M. Vloothuis<sup>2</sup>, Maayken van den Berg<sup>3</sup>, Maria Crotty<sup>3</sup>, Gert Kwakkel<sup>1,2,4,5</sup>, Erwin E. H. van Wegen<sup>1,4</sup>

**1** Department of Rehabilitation Medicine, Amsterdam Movement Sciences, Amsterdam UMC, VU University Medical Centre, Amsterdam, The Netherlands, **2** Rehabilitation Research Centre, Reade, Amsterdam, The Netherlands, **3** Department of Rehabilitation, Aged and Extended Care, Flinders University, Adelaide, Australia, **4** Amsterdam Neuroscience, Amsterdam, The Netherlands, **5** Department of Physical Therapy and Human Movement Sciences, Northwestern University, Chicago, Illinois, United States of America

\* [r.nijland@reade.nl](mailto:r.nijland@reade.nl)



## OPEN ACCESS

**Citation:** Mulder M, Nijland RHM, Vloothuis JDM, van den Berg M, Crotty M, Kwakkel G, et al. (2022) Comparing two identically protocolized, multicentre, randomized controlled trials on caregiver-mediated exercises poststroke: Any differences across countries? PLoS ONE 17(1): e0263013. <https://doi.org/10.1371/journal.pone.0263013>

**Editor:** Abiodun E. Akinwuntan, University of Kansas Medical Center, UNITED STATES

**Received:** July 6, 2021

**Accepted:** January 10, 2022

**Published:** January 25, 2022

**Copyright:** © 2022 Mulder et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its [Supporting information](#) files.

**Funding:** This study was financially supported by the Netherlands Organisation for Health Research and Development (ZonMw – [www.zonmw.nl](http://www.zonmw.nl)) grant number: 837001408 (2013) and 630000002 (2015), and ZonMw / Scientific College Physiotherapy (WCF – [www.kngf.nl](http://www.kngf.nl)) grant number:

## Abstract

### Background

The evidence for rehabilitation interventions poststroke lack sufficient robustness. However, variation in treatment effects across countries have been given little attention.

### Objective

To compare two identically protocolized trials conducted in different western countries in order to identify factors that may have caused variation in secondary trial outcomes.

### Methods

Comparative study based on individual patient data (N = 129) from two randomized controlled trials, conducted in hospitals and rehabilitation facilities in the Netherlands (N = 66) and Australia (N = 63). Patients with stroke and their caregivers were randomly allocated to an 8-week caregiver-mediated exercises intervention (N = 63; 31 Australian and 32 Dutch) or to a control group (N = 66; 32 Australian and 34 Dutch). Patient characteristics, compliance, usual care and process measures were compared across countries. We examined if study setting significantly moderated the trial outcomes: Hospital Anxiety and Depression Scale, Fatigue Severity Scale and General Self-Efficacy Scale, measured at 8- and 12 weeks follow-up. In addition, we explored if factors that were significantly different across countries caused variation in these trial outcomes.

### Results

Most patients suffered an ischemic stroke, were in the subacute phase and participated with their partner. Dutch patients were younger ( $P = 0.005$ ) and had a lower functional status ( $P = 0.001$ ). Australian patients were recruited earlier poststroke ( $P < 0.001$ ), spent less time in

858001102 (2018). Grants were awarded to GK and EvW. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

exercise therapy ( $P < 0.001$ ) and had a shorter length of stay ( $P < 0.001$ ). The level of contamination was higher ( $P = 0.040$ ) among Dutch controls. No effect modification was observed and trial outcomes did not change after controlling for cross-country differences.

## Conclusions

The present study highlighted important clinical differences across countries whilst using an identical study protocol. The observed differences could result in a different potential for recovery and variation in treatment effects across trials. We argue that we can proceed faster to evaluating interventions within international pragmatic trials.

## Introduction

Stroke affects approximately 104 million people worldwide and the global incidence of stroke is currently estimated at 11.9 million [1]. Stroke is a disabling condition with variable recovery patterns and a heterogeneous functional outcome [2]. Most patients require rehabilitation to assist functional recovery and optimize independence in daily life [2, 3]. However, the current evidence for specific rehabilitation interventions poststroke lacks sufficient robustness and is largely based on findings from small, phase II, Randomized Controlled Trials (RCTs) [3–5]. Phase II RCTs are designed to estimate treatment effects with the lowest possible bias. Therefore, conflicting results across trials is assumed to result from either imprecision of effect estimates or true effect variation as a result of clinical heterogeneity [6]. Clinical heterogeneity refers to systematic differences in the outcomes studied or timing of outcome measurement, characteristics of patients (e.g. age, ethnicity, disease severity), interventions (e.g. content, dose, duration), or study setting (e.g. country, healthcare system, cultural context) [7]. Further investigation of responder and non-responder groups or dose-response relationships is common within stroke recovery trials. However, variation in treatment effects across countries with a different healthcare system and cultural context have been given little attention.

In a recent collaboration, we simultaneously conducted two RCT's in Australia (ACTRN12613000779774) and the Netherlands (NTR4300). Both studies investigated the effect of a Caregiver-Mediated Exercises (CME) intervention with e-health support [8, 9], which may be a valuable intervention to augment exercise therapy and improve functional outcome after stroke [10, 11]. Although the primary outcome of both trials was neutral with respect to patients' self-reported mobility, different significant results regarding secondary psychosocial outcomes were found. The Dutch trial reported a significant decrease in caregiver depression and patient anxiety [9], while the Australian trial reported a significant decrease in caregiver fatigue and increased caregiver self-efficacy [8]. Both trials had an identical study design that obeys CONSORT statements, used the same measurements of outcome and identical criteria of patient selection [12]. In addition, an identical intervention protocol in terms of content, dosing and treatment goals was used [13]. However, the studies were embedded within a different healthcare system and cultural context, making study setting the main source of heterogeneity for possible differences in outcome. Combining individual patient data from both trials allows us to investigate clinical heterogeneity and true effect variation across countries providing valuable information for future trial design and conduct.

The aim of the current study was to compare two identically protocolized trials in terms of design, outcome measurement, criteria of patient selection, and intervention protocol, applied in the Netherlands and Australia in order to identify factors that may have caused variation

across countries in secondary trial outcomes. First, we hypothesized that the identical criteria of patient selection reduced patient heterogeneity. Second, we hypothesized that significant differences exist across countries in terms of usual exercise therapy and processes of care including length of inpatient stay and timing of recruitment. Finally, we hypothesized that study setting significantly moderated the secondary trial outcomes anxiety, depression, fatigue and self-efficacy.

## Material and methods

### Study design

This study combined individual patient data from two identically protocolized, observer-blinded, multicentre RCTs in which patients with stroke and their appointed caregivers were randomly allocated to a CME intervention with e-health support in addition to usual care or to a control group that received usual care alone [12]. The two trials were simultaneously started in the Netherlands and Australia, respectively. The randomization procedures, sample size calculations and results of the separate trials have been published previously [8, 9]. The medical ethics committee of the Slotervaart Hospital and Reade approved the Dutch study (NTR4300) and the local health research ethics board of Southern Adelaide Research Ethics Committee approved the Australian study (ACTRN12613000779774).

### Definitions

Imprecision is defined as random error in trial outcomes, as a result of sample variation. Hence, precision of effect estimates depends on the sample size, with smaller studies being less precise [6].

Methodological heterogeneity is defined as a difference across trials in the study design, outcome measurement tools or risk of bias. A higher degree of bias will lead to systematic error in trial outcomes [6].

Clinical heterogeneity is defined as a systematic difference across trials in the outcomes studied or timing of outcome measurement, characteristics of patients (e.g. age, ethnicity, disease severity), interventions (e.g. content, dose, duration), or study setting (e.g. country, healthcare system, cultural context) [6, 7].

True effect variation is defined as variation in treatment effects as a consequence of clinical heterogeneity. True variation in trial outcomes is not explained by random or systematic error [6, 7].

Study protocol is defined as a combination of the study design and intervention protocol. The design characteristics (e.g., blinding, randomization), study procedures, outcome measures, power calculation, criteria of patient selection and statistical analysis are protocolized in the study design. The essential elements of the trial intervention including its content, mode of delivery, dose, duration and monitoring of compliance are protocolized in the intervention protocol.

Study setting is defined as the geographical location of a trial (i.e., country) including the healthcare system and cultural context in which it is embedded, and the context of individual study sites including their level of care (i.e., primary, secondary, tertiary), local healthcare processes and available resources including staff and expertise [14].

Contamination is defined as the receipt of active intervention amongst participants in the control arm of an RCT [15].

Pragmatic trials are defined as RCTs designed to determine the effect of an intervention under the usual conditions in which it will be applied [16], more specifically the trial should 1)

enrol a real-world population; 2) be conducted in a real-world setting; 3) include an appropriate comparison arm; and 3) capture relevant outcomes [17].

### Participant recruitment

Patients in the Netherlands were recruited from four hospital stroke units, two rehabilitation centres and seven rehabilitation wards of nursing homes in Amsterdam and its near surroundings. Patients in Australia were recruited from two hospital stroke units and one hospital rehabilitation unit in metropolitan Adelaide. Eligibility criteria for both patients and caregivers were: 1) 18 years or older; 2) able to understand the Dutch or English language; and 3) no significant signs of depression (Hospital Anxiety and Depression Scale [HADS] depression subscale < 11). Patients were eligible if they: 1) were diagnosed with stroke according to the WHO definition [18]; 2) experienced mobility limitations (Functional Ambulation Categories score < 5); 3) had sufficient cognition to follow instructions and provide informed consent (Mini-Mental State Examination [MMSE] > 18); 4) lived independently prior to stroke; 5) were planned to be discharged home; and 6) were willing and able to appoint a caregiver. Caregivers were eligible if they were: 1) willing to participate in CME; and 2) physically able to support the patient. Exclusion criteria for both patients and caregivers were: 1) a serious comorbidity that interfered with participation; and 2) not being medically stable. Participants were recruited in the early rehabilitation phase, following admission in a participating centre. All participants provided written informed consent prior to study enrolment.

### Intervention

The C4S program consisted of an 8-week incremental training program with task-oriented and mobility-related exercises. An identical treatment protocol, including the intended dose, following the TIDieR guidelines was used [13]. Patient-caregiver dyads were asked to perform exercises together, at least five times a week for 30 minutes, within the clinical setting, outside or at home. In addition, participants were encouraged to exercise outside usual training hours. During a weekly face-to-face session with a trained physical therapist, dyads were instructed, evaluated and caregivers were educated to become an exercise coach. During the program, dyads were supported by an offline e-health application with instruction videos, in addition to usual care. The Australian and Dutch trials used respectively an English and Dutch version of the same video application.

### Primary and secondary outcomes

Patients and caregivers in both trials were assessed using the same outcome measures, at baseline prior to randomization, and at 8- and 12-weeks follow-up. In both countries, outcome assessments were completed by an independent assessor blinded to treatment allocation. The primary outcome measure of both trials was the mobility domain of the Stroke Impact Scale (SIS 3.0) [19]. The secondary outcome measures including their description and references has been published previously [12]. For the purpose of the current investigation, we selected the anxiety and depression subscales of the HADS [20], the Fatigue Severity Scale (FSS) [21] and General Self-Efficacy Scale (GSES) [22] of both patients and caregivers to assess secondary trial outcomes of anxiety, depression, fatigue and self-efficacy.

To investigate systematic differences across countries in patient characteristics we used sex, age, living situation prior to stroke (alone or with someone), relationship with the appointed caregiver, type of stroke (ischemic or haemorrhagic), site of stroke (left or right hemisphere), functional status measured with the modified Rankin Scale (mRS) [23] and cognition measured with the MMSE [24] at the moment of recruitment. Self-reported exercise minutes from

diaries that were kept for 8 weeks following randomization, were used to investigate systematic differences in compliance and contamination (CME minutes) and usual exercise therapy (exercise minutes with a nurse, therapist and independent exercise minutes). Process measures that were compared across countries included timing of recruitment (days between stroke onset and randomization), timing of discharge (days between randomization and inpatient discharge) and Length Of inpatient Stay (LOS) which was defined as the number of days between hospital admission and discharge from inpatient care.

### Statistical analyses

Individual patient data from both trials were merged in one pooled database. First, systematic differences across the Dutch and Australian trials in terms of 1) patient characteristics; 2) compliance and usual exercise therapy; and 3) process measures; were compared using independent samples t-tests and we reported means and standard deviations in case of a normal distribution or used Mann-Whitney U tests and reported medians and interquartile ranges in case of a skewed distribution. A Chi-square test was used for categorical outcomes or a Fisher's exact test in case of few observations for individual cells (<10). Second, Linear Mixed Models were used to examine if the dichotomous variable 'country' was an effect-modifier that significantly moderated the secondary trial outcomes, with HADS anxiety, HADS depression, FSS and GSES as the dependent variables and treatment allocation, country, baseline score of the dependent variable and a treatment-by-country interaction term as covariates. A random intercept was included to adjust for the dependency of the repeated observations within subjects [25]. If a significant interaction was detected, treatment effects were reported separately for both countries. If no significant interaction was identified, the overall treatment effect was reported while controlling for country. Finally, patient characteristics and process measures that were systematically different across countries were successively added to the longitudinal model in order to explore if these factors may have caused variation in secondary trial outcomes. If the regression coefficient of treatment allocation changed by  $\geq 10\%$ , the covariate was considered to be a confounder and new candidate confounders were added to the longitudinal model [26, 27]. A 2-tailed significance level alpha of 0.05 was used for all the statistical tests that were performed. Data were analysed using IBM SPSS Statistics version 25 for Windows.

### Results

Individual patient data of 129 patients and their appointed caregivers were included, of which 66 dyads were recruited in the Netherlands and 63 dyads in Australia. A total of 617/709 screened (87.0%) in Australia and 960/1082 screened (88.7%) in the Netherlands were excluded from participation. Patients were selected according to the same eligibility criteria, and no additional strategies were used to select or exclude patients. No significant difference ( $P = 0.054$ ) was observed in the proportion of eligible patient-caregiver dyads that declined to participate in Australia, 29/709 screened (4.1%), compared to those that declined to participate in the Netherlands, 56/1082 screened (5.2%).

### Systematic differences in patient characteristics

Table 1 presents the baseline characteristics of recruited patients in both countries. Patients in the Dutch trial were significantly younger (MD 7.7 years, CI 2.3–13.0;  $P = 0.005$ ) and had a lower functional level (82% with a moderate to severe disability versus 54% in Australia) when compared to Australian patients. Although a higher proportion of patients were living alone prior to stroke in the Netherlands, this difference did not reach significance.

**Table 1. Differences in patient characteristics and process measures.**

Patient characteristics	The Netherlands (n = 66)	Australia (n = 63)	P-value
Age in years, mean (SD)	59.9 (14.8)	67.5 (15.8)	<b>0.005</b> <sup>a</sup>
Sex, number of males (%)	41 (62.1)	40 (63.5)	0.872 <sup>c</sup>
Type of stroke, N (%)			0.115 <sup>d</sup>
Ischemic	50 (75.8)	54 (85.7)	
Haemorrhagic	16 (24.2)	8 (12.7)	
Both	0 (0.0)	1 (1.6)	
Site of stroke, N (%)			0.592 <sup>d</sup>
Right hemisphere	37 (56.1)	39 (61.9)	
Left hemisphere	28 (42.4)	24 (38.1)	
Brainstem	1 (1.5)	0 (0.0)	
mRS disability (0–5), N (%)			<b>0.001</b> <sup>a,d</sup>
1 no significant	1 (1.5)	2 (3.2)	
2–3 slight-moderate	11 (16.7)	27 (42.9)	
4–5 moderate-severe	54 (81.8)	34 (53.9)	
MMSE (0–30), median (IQR)	28.0 (25.0–29.0)	26.0 (24.0–28.3)	0.226 <sup>b</sup>
Living alone, N (%)	20 (30.3)	10 (15.9)	0.059 <sup>c</sup>
Caregiver, N (%)			0.280 <sup>c</sup>
Partner	39 (59.1)	43 (68.3)	
Another caregiver	27 (40.9)	20 (31.7)	
Timing of recruitment in days, median (IQR)	37 (28–56)	14 (7–22)	< <b>0.001</b> <sup>a,b</sup>
Timing of discharge in days, median (IQR)	69 (41–92)	18 (7–35)	< <b>0.001</b> <sup>a,b</sup>
LOS in days, median (IQR)	112 (79–148)	34 (18–56)	< <b>0.001</b> <sup>a,b</sup>

IQR, inter quartile range; SD, standard deviation; mRS, modified Rankin Scale; MMSE, Mini-Mental State Examination; LOS, length of inpatient stay.

<sup>a</sup> *p* values were calculated using independent samples t-tests,

<sup>b</sup> Mann-Whitney U tests,

<sup>c</sup> Chi-square tests,

<sup>d</sup> Fisher's exact tests

<https://doi.org/10.1371/journal.pone.0263013.t001>

## Systematic differences in compliance and usual care

Compliance with the intervention protocol did not differ across Australian (N = 31, Median 953 CME minutes, IQR 635–1390) and Dutch (N = 32, Median 1150 CME minutes, IQR 850–1500) intervention groups (P = 0.383). However, Australian controls reported significantly fewer CME minutes (N = 32, Median 60 minutes, IQR 5–495; P = 0.040) than Dutch controls (N = 34, Median 350 minutes; IQR 95–1065). Patients in Australia reported significantly (P = 0.04) less usual exercise minutes during the 8 weeks following randomization (N = 63, Median 2550 minutes, IQR 1478–3683) when compared to the Netherlands (N = 66, Median 3403 minutes, IQR 2220–3900). More specifically, patients in Australia reported less minutes in formal exercise therapy (Median 1080 minutes, IQR 473–1735; versus Median 2025 minutes, IQR 1368–2858; P < 0.001), but more independent exercises (Median 930 minutes, IQR 300–2408; versus Median 550 minutes, IQR 190–1200; P < 0.037). We found no significant difference (P = 0.114) in exercise time with a nurse (Median 0 minutes, IQR 0–140 and Median 35 minutes, IQR 0–173, in Australia and the Netherlands respectively).

## Systematic differences in process measures

In [Table 1](#) process measures are presented for both countries. All process measures differed significantly across countries (P < 0.001). Patient-caregiver dyads in the Australian trial were



Table 2. Secondary treatment effects and study setting interaction effects.

Outcome measures patient	Beta (95% CI)	P-value <sup>a</sup> Treatment	P-value <sup>a</sup> Treatment*Country
HADS anxiety (0–21)	1.10 (0.01–2.20)	<b>0.049*</b>	0.220
HADS depression (0–21)	0.66 (-0.40–1.72)	0.222	0.470
FSS (1–7)	0.19 (-0.31–0.68)	0.462	0.505
GSES (10–40)	-0.42 (-1.99–1.14)	0.593	0.942
Outcome measures caregiver	Beta (95% CI)	P-value <sup>a</sup> Treatment	P-value <sup>a</sup> Treatment*Country
HADS anxiety (0–21)	0.23 (-0.94–1.41)	0.692	0.541
HADS depression (0–21)	0.94 (0.00–1.88)	<b>0.050*</b>	0.412
FSS (1–7)	0.18 (-0.19–0.54)	0.341	0.990
GSES (10–40)	-0.44 (-1.71–0.83)	0.493	0.703

HADS, Hospital Anxiety and Depression Scale; FSS, Fatigue Severity Scale; GSES, General Self-Efficacy Scale.

<sup>a</sup> *p* values were calculated using Linear Mixed Model analyses

<https://doi.org/10.1371/journal.pone.0263013.t002>

recruited significantly earlier poststroke when compared to the Netherlands. Moreover, LOS was significantly shorter in Australia and as a result discharge occurred earlier in the program. Australian patients generally returned home in the first five weeks following randomization while Dutch patients returned home in the last two weeks or after completing the 8-week program.

### Differences in secondary trial outcomes across countries

No significant interaction effects were observed between treatment allocation and the variable ‘country’ with respect to anxiety, depression, self-efficacy and fatigue of both patients and caregivers. The overall treatment effects of the C4S program with country as a covariate, and the interaction effects of treatment allocation and country are summarized in Table 2. The observed trial outcomes did not change after controlling for age, baseline functional status, or after controlling for usual exercise dose, timing of recruitment, timing of discharge and LOS.

### Discussion

To our knowledge, this is the first study in stroke rehabilitation that investigates cross-cultural clinical differences by comparing two identically protocolized phase II trials with respect to design, outcome measurement, patient selection and intervention conducted in two different Western countries. The present study shows that, despite identical criteria of patient selection, systematic clinical differences may be observed in study participants across countries. In addition, heterogeneity was observed in the level of contamination, amount of usual care and total length of inpatient stay. Awareness of clinical heterogeneity across countries is especially important when designing and conducting global trials. However, study setting did not moderate the secondary trial outcomes anxiety, depression, fatigue and self-efficacy suggesting that we can proceed faster to evaluating interventions within heterogeneous populations and study settings in pragmatic phase III and IV trials.

Patient heterogeneity was observed across countries with older age and a higher functional level at baseline in Australia when compared to the Netherlands. In general, age and functional status after stroke are factors known to influence discharge disposition and referral to inpatient rehabilitation after acute care on a Hospital Stroke Unit (HSU) [28–30]. In the Dutch study, participating HSUs were unable to recruit patients as a result of their fast discharge policies with a short LOS in the hospital. Hence, all patients were recruited from rehabilitation centres

and rehabilitation wards of nursing homes. In contrast, Australian patients were recruited from acute HSUs as well as from Hospital Rehabilitation Units which may explain the different casemix. In parallel, Australian patients were recruited earlier poststroke which is not unexpected within an acute setting. Age and functional status at admission are significant predictors of functional outcome after stroke [28, 31]. Moreover, functional improvement is significantly associated with progress of time [32]. This suggests that age, functional status and time poststroke at recruitment may be important effect modifiers that can cause variation in functional outcomes across trials despite following an identical study protocol.

The present study also shows that usual care is strongly dependent on features of the health-care system and cultural context. To our knowledge, there are no recent studies reporting about practice variation in usual care across countries. We show that Australian patients spent less time in formal exercise therapy in both groups and total LOS was shorter when compared to the Netherlands. For the purpose of this study, LOS was defined as the number of days between acute hospital admission and final discharge home. Hence, this finding is not explained by timing of recruitment or level of care. The shorter LOS in Australia may be explained by the less severely affected population [33, 34] or reflect differences in local health-care policies with respect to rehabilitation services such as Early Supported Discharge (ESD) [35]. These findings are supported by lower costs of care observed in Australia, when compared to the Netherlands [36]. Although the emphasis on ESD has accelerated the interest in outpatient services [37], most resources in stroke rehabilitation are focused on inpatient care [35, 38]. As a result of the shorter LOS, Australian patients returned home faster following randomization which could explain the lower dose of formal exercise therapy during the 8 weeks thereafter. In stroke rehabilitation trials, the control group often receives 'usual care'. Unfortunately, the content, timing and dose of usual care is often poorly described [39]. Acknowledging that a difference in the amount of usual care delivered can result in a different potential for recovery [40, 41] it is important to realize that these differences may exist across countries when designing global trials.

In the two trials, it was intended that controls received only usual care [12, 13]. Although some contamination was present in the control groups of both studies, a significant difference was observed in its magnitude. The self-reported dose of CME within control groups was about 8 weekly minutes among Australian controls and 44 weekly minutes among Dutch controls, resulting in a larger treatment contrast in Australia. One of the most common sources of contamination within clinical trials is communication between trial arms, either between staff members or participants [15]. Such communication is more likely within an inpatient environment where patients are admitted to the same rehabilitation ward and providers of the intervention work closely together with other healthcare professionals. The earlier timing of discharge in Australia may therefore partly explain the lower level of contamination in this study. Another common source of contamination is when the intervention is already part of usual care to some extent [15]. Although the involvement of caregivers may differ across different social, cultural, political and organizational contexts [10], the total level of contamination is low and we have no data to support the structural use of CME in a usual care context in the Netherlands.

The present study highlighted important clinical differences across countries whilst using an identical study protocol. Although different significant results were reported in the separate trials regarding secondary outcomes of anxiety, depression, self-efficacy and fatigue [8, 9], we did not find a moderating effect of study setting on these trial outcomes. This suggests that treatment effects on secondary outcomes within small phase II trials may be imprecise and should be interpreted with caution. We argue that we should proceed faster to evaluating promising rehabilitation interventions poststroke in large phase III and IV trials. International



collaboration to standardize study protocols including study design, outcome measurement and reporting is an important next step to improve robustness of trial outcomes in the field of stroke rehabilitation [42–44]. Awareness of clinical heterogeneity and consensus in trial conduct across countries through the international GAINS network is especially important when designing future global trials. First, we recommend that patients are recruited based on broad selection criteria, however at fixed times or within a narrow time window poststroke, and from multiple study settings to improve generalizability. Subsequently, responder analyses can be used to investigate and interpret a heterogeneous response to treatment. Second, the control condition should be monitored during the study period and we should improve our reporting of ‘usual care’ including its specific content, dose and location. Third, careful consideration of the processes driving contamination is needed. Although cluster randomization is often mentioned as a solution, it could lead to patient heterogeneity across trial arms when patients are recruited from different study settings [45]. A balanced distribution of clusters over different settings or changes in trial conduct to reduce contamination could be considered. Finally, international consensus on parameters of trial conduct and reporting is essential for improving the quality of future stroke rehabilitation trials.

### Strengths and limitations

Our individual patient dataset from two identically protocolized trials, independently conducted in two countries with a different healthcare system and cultural context, is unique. However, several study limitations should be acknowledged. First, the current study investigated differences across two Western, high-income countries and no generalizations can therefore be made to non-Western, low- or middle-income countries. Second, we did not collect information on ethnicity, racial groups or comorbidities. Third, timing of recruitment was standardized based on the moment of admission to a participating rehabilitation ward. As a result, randomization, outcome measurement and initiation of the C4S program occurred at relatively arbitrary time points poststroke.

### Conclusion

The present study highlighted important clinical differences across countries whilst using an identical study protocol, providing valuable information for future trial design and conduct. Theoretically, the observed clinical differences could result in a different potential for recovery and variation in treatment effects across trials. However, we did not find a moderating effect of study setting on the current secondary trial outcomes. We argue that we can proceed faster to evaluating promising interventions within large, pragmatic trials and give recommendations for designing future stroke rehabilitation trials.

### Supporting information

**S1 Dataset.**  
(SAV)

### Acknowledgments

The Care4Stroke project was conducted by M. van den Berg (AUS), M. Crotty (AUS), Q. Goedhart (NL), G. Kwakkel (NL), M. Mulder (NL), R. Nijland (NL), J. Vloothuis (NL), and E. van Wegen (NL). We would like to thank the Australian and Dutch patient-caregiver dyads that participated in the Care4Stroke project and physical therapists and physicians from the participating study sites.

## Author Contributions

**Conceptualization:** Marijn Mulder, Rinske H. M. Nijland, Judith D. M. Vloothuis, Gert Kwakkel, Erwin E. H. van Wegen.

**Formal analysis:** Marijn Mulder.

**Funding acquisition:** Rinske H. M. Nijland, Gert Kwakkel, Erwin E. H. van Wegen.

**Investigation:** Marijn Mulder, Rinske H. M. Nijland, Judith D. M. Vloothuis, Maayken van den Berg, Maria Crotty.

**Methodology:** Marijn Mulder, Rinske H. M. Nijland, Gert Kwakkel, Erwin E. H. van Wegen.

**Project administration:** Marijn Mulder, Rinske H. M. Nijland, Maayken van den Berg, Maria Crotty, Gert Kwakkel, Erwin E. H. van Wegen.

**Supervision:** Rinske H. M. Nijland, Gert Kwakkel, Erwin E. H. van Wegen.

**Writing – original draft:** Marijn Mulder.

**Writing – review & editing:** Rinske H. M. Nijland, Judith D. M. Vloothuis, Maayken van den Berg, Maria Crotty, Gert Kwakkel, Erwin E. H. van Wegen.

## References

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018; 392:1789–858. [https://doi.org/10.1016/S0140-6736\(18\)32279-7](https://doi.org/10.1016/S0140-6736(18)32279-7) PMID: 30496104
2. Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. *Lancet*. 2011; 377:1693–702. [https://doi.org/10.1016/S0140-6736\(11\)60325-5](https://doi.org/10.1016/S0140-6736(11)60325-5) PMID: 21571152
3. Stinear CM, Lang CE, Zeiler S, Byblow WD. Advances and challenges in stroke rehabilitation. *Lancet Neurol*. 2020; 19(4):348–60. [https://doi.org/10.1016/S1474-4422\(19\)30415-6](https://doi.org/10.1016/S1474-4422(19)30415-6) PMID: 32004440
4. Veerbeek JM, van Wegen EEH, Van Peppen R, Van Der Wees P, Hendriks E, Rietberg M, et al. What is the evidence for physical therapy poststroke? A systematic review and meta-analysis. *PLoS One*. 2014; 9(2). <https://doi.org/10.1371/journal.pone.0087987> PMID: 24505342
5. Winstein C. The ATTEND trial: An alternative explanation with implications for future recovery and rehabilitation clinical trials. *Int J Stroke*. 2018; 13(2):112–6. <https://doi.org/10.1177/1747493017743061> PMID: 29214907
6. Higgins J, Green S. *Cochrane handbook for systematic reviews of interventions*. Cochrane Database Syst Rev. 2008;187–235. <https://doi.org/https%3A//doi.org/10.1002/9780470712184>
7. Gagnier J, Morgenstern H, Altman D, Berlin J, Chang S, McCulloch P, et al. Consensus-based recommendations for investigating clinical heterogeneity in systematic reviews. *BMC Med Res Methodol*. 2013; 13(1): 1–11. <https://doi.org/10.1186/1471-2288-13-106> PMID: 24004523
8. van den Berg M, Crotty M, Liu E, Killington M, Kwakkel G, van Wegen EEH. Early supported discharge by caregiver-mediated exercises and e-health support after stroke: A proof-of-concept trial. *Stroke*. 2016; 47(7):1885–92. <https://doi.org/10.1161/STROKEAHA.116.013431> PMID: 27301941
9. Vloothuis, Mulder M, Nijland RHM, Konijnenbelt M, Mulder H, Hertogh CMPM, et al. Caregiver-mediated exercises with e-health support for early supported discharge after stroke (CARE4STROKE): A randomized controlled trial. *PLoS One*. 2019; 14(4):e0214241. <https://doi.org/10.1371/journal.pone.0214241> PMID: 30958833
10. Vloothuis JDM, Mulder M, Veerbeek JM, Konijnenbelt M, Visser-Meily JMA, Ket JCF. Caregiver-mediated exercises for improving outcomes after stroke. *Cochrane Database Syst Rev*. 2016; 12:1465–858. <https://doi.org/10.1002/14651858.CD011058.pub2> PMID: 28002636
11. Yasmeen I, Krewulak KD, Grant C, Stelfox HT, Fiest KM. The Effect of Caregiver-Mediated Mobility Interventions in Hospitalized Patients on Patient, Caregiver, and Health System Outcomes: A Systematic Review. *Arch Rehabil Res Clin Transl*. 2020; 2(3):100053. <https://doi.org/10.1016/j.arrct.2020.100053> PMID: 33543080
12. Vloothuis JDM, Mulder M, Nijland RHM, Konijnenbelt M, Mulder H, Hertogh CMPM, et al. Caregiver-mediated exercises with e-health support for early supported discharge after stroke (CARE4STROKE):

- study protocol for a randomized controlled trial. *BMC Neurol.* 2015; 15(193). <https://doi.org/10.1186/s12883-015-0440-z> PMID: 26452543
13. Vloothuis JDM, de Bruin J, Mulder M, Nijland RHM, Kwakkel G, van Wegen EEH. Description of the CARE4STROKE programme: A caregiver-mediated exercises intervention with e-health support for stroke patients. *Physiother Res Int.* 2018; 23(3):e1719. <https://doi.org/10.1002/pri.1719> PMID: 29797740
  14. Rothwell PM. External validity of randomised controlled trials: "To whom do the results of this trial apply?" *Lancet.* 2005; 365(9453):82–93. [https://doi.org/10.1016/S0140-6736\(04\)17670-8](https://doi.org/10.1016/S0140-6736(04)17670-8) PMID: 15639683
  15. Magill N, Knight R, McCrone P, Ismail K, Landau S. A scoping review of the problems and solutions associated with contamination in trials of complex interventions in mental health. *BMC Med Res Methodol.* 2019; 19(4):1–13. <https://doi.org/10.1186/s12874-018-0646-z> PMID: 30616508
  16. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): A tool to help trial designers. *C Can Med Assoc J.* 2009; 62(5):464–75. <https://doi.org/https%3A/doi.org/10.1016/j.jclinepi.2008.12.011> PMID: 19348971
  17. Gamerman V, Cai T, Elsäßer A. Pragmatic randomized clinical trials: best practices and statistical guidance. *Heal Serv Outcomes Res Methodol.* 2019; 19(1):23–35. <https://doi.org/https%3A/doi.org/10.1007/s10742-018-0192-5>
  18. Aho K, Harmsen P, Hatano S, Marquardsen J, Smirnov VE, Strasser T. Cerebrovascular disease in the community: results of a WHO collaborative study. *Bull World Health Organ.* 1980; 58(1):113–30. PMID: 6966542.
  19. Duncan PW, Bode RK, Lai SM, Perera S. Rasch analysis of a new stroke-specific outcome scale: The stroke impact scale. *Arch Phys Med Rehabil.* 2003; 84(7):950–63. [https://doi.org/10.1016/s0003-9993\(03\)00035-2](https://doi.org/10.1016/s0003-9993(03)00035-2) PMID: 12881816
  20. Bjelland I, Dahl A, Haug T, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *J Psychosom Res.* 2002; 52(2):69–77. [https://doi.org/10.1016/s0022-3999\(01\)00296-3](https://doi.org/10.1016/s0022-3999(01)00296-3) PMID: 11832252
  21. Valko P, Bassetti C, Bloch K, Held U, Baumann C. Validation of the fatigue severity scale in a Swiss cohort. *Sleep.* 2008; 31(11):1601–7. <https://doi.org/10.1093/sleep/31.11.1601> PMID: 19014080
  22. Schwarzer R, Jerusalem M. Generalized Self-Efficacy scale. In: Wright JW, S JM, editors. *Measures in health psychology: a user's portfolio Causal and control beliefs.* Windsor, UK: NFER-NELSON; 1995. p. 35–7.
  23. de Haan R, Limburg M, Bossuyt P, van der Meulen J, Aaronson N. The clinical meaning of Rankin "handicap" grades after stroke. *Stroke.* 1995; 26(11):2027–30. <https://doi.org/10.1161/01.str.26.11.2027> PMID: 7482643
  24. Tombaugh T, McIntyre N. The mini-mental state examination: a comprehensive review. *J Am Geriatr Soc.* 1992; 40(9):922–35. <https://doi.org/10.1111/j.1532-5415.1992.tb01992.x> PMID: 1512391
  25. Twisk JW. *Applied Longitudinal Data Analysis for Epidemiology.* 2nd ed. Cambridge; 2013.
  26. Twisk JWR. *Inleiding in de toegepaste biostatistiek. Inleiding in de toegepaste biostatistiek.* 2016. Dutch.
  27. Mickey R, Greenland S. The impact of confounder selection criteria on effect estimation. *Am J Epidemiol.* 1989; 129(1):125–137. <https://doi.org/10.1093/oxfordjournals.aje.a115101> PMID: 2910056
  28. Hakkennes SJ, Brock K, Hill KD. Selection for inpatient rehabilitation after acute stroke: A systematic review of the literature. *Arch Phys Med Rehabil.* 2011; 92(12):2057–70. <https://doi.org/10.1016/j.apmr.2011.07.189> PMID: 22133256
  29. Dutrieux RD, Van Eijk M, Van Mierlo ML, Van Heugten CM, Visser-Meily JMA, Achterberg WP. Discharge home after acute stroke: Differences between older and younger patients. *J Rehabil Med.* 2016; 48(1):14–8. <https://doi.org/10.2340/16501977-2045> PMID: 26667264
  30. Labberton AS, Barra M, Rønning OM, Thommessen B, Churilov L, Cadilhac DA, et al. Patient and service factors associated with referral and admission to inpatient rehabilitation after the acute phase of stroke in Australia and Norway. *BMC Health Serv Res.* 2019; 19(1):1–10.
  31. Fahey M, Crayton E, Wolfe C, Douiri A. Clinical prediction models for mortality and functional outcome following ischemic stroke: A systematic review and meta-analysis. *PLoS One.* 2018; 13(1):e0185402. <https://doi.org/10.1371/journal.pone.0185402> PMID: 29377923
  32. Kwakkel G, Kollen B, Twisk J. Impact of time on improvement of outcome after stroke. *Stroke.* 2006; 37(9):2348–53. <https://doi.org/10.1161/01.STR.0000238594.91938.1e> PMID: 16931787
  33. Grant C, Goldsmith CH, Anton HA. Inpatient stroke rehabilitation lengths of stay in Canada derived from the national rehabilitation reporting system, 2008 and 2009. *Arch Phys Med Rehabil.* 2014; 95(1):74–8. <https://doi.org/10.1016/j.apmr.2013.08.014> PMID: 24001444

34. Lugo-Palacios DG, Gannon B, Gittins M, Vail A, Bowen A, Tyson S. Variations in hospital resource use across stroke care teams in England, Wales and Northern Ireland: a retrospective observational study. *BMJ Open*. 2019; 9(9):e030426. <https://doi.org/10.1136/bmjopen-2019-030426> PMID: 31542751
35. Langhorne P, Baylan S. Early supported discharge services for people with acute stroke. *Cochrane Database Syst Rev*. 2017; 7:1465–858. <https://doi.org/10.1002/14651858.CD000443.pub4> PMID: 28703869
36. Rajsic S, Gothe H, Borba HH, Sroczynski G, Vujicic J, Toell T, et al. Economic burden of stroke: a systematic review on post-stroke care. *Eur J Health Econ*. 2019; 20:107–34. <https://doi.org/10.1007/s10198-018-0984-0> PMID: 29909569
37. Outpatient Service Trialists. Therapy-based rehabilitation services for stroke patients at home. *Cochrane Database Syst Rev*. 2003; 1:1465–858. <https://doi.org/10.1002/14651858.CD002925> PMID: 12535444
38. Luengo-Fernandez R, Violato M, Candio P, Leal J. Economic burden of stroke across Europe: A population-based cost analysis. *Eur Stroke J*. 2020; 5(1):17–25. <https://doi.org/10.1177/2396987319883160> PMID: 32232166
39. Lohse KR, Pathania A, Wegman R, Boyd LA, Lang CE. On the reporting of experimental and control therapies in stroke rehabilitation trials: a systematic review. *Arch Phys Med Rehabil*. 2018; 99(7):1424–32. <https://doi.org/10.1016/j.apmr.2017.12.024> PMID: 29412168
40. Lohse KR, Lang CE, Boyd LA. Is more better? Using metadata to explore dose–response relationships in stroke rehabilitation. *Stroke*. 2014; 45:2053–8. <https://doi.org/10.1161/STROKEAHA.114.004695> PMID: 24867924
41. Winstein C, Kim B, Kim S, Martinez C, Schweighofer N. Dosage matters. A phase IIb randomized controlled trial of motor therapy in the chronic phase after stroke. *Stroke*. 2019; 50:1831–7. <https://doi.org/10.1161/STROKEAHA.118.023603>
42. Bernhardt J, Borschmann K, Boyd L, Carmichael ST, Corbett D, Cramer SC, et al. Moving rehabilitation research forward: Developing consensus statements for rehabilitation and recovery research. *Int J Stroke*. 2016; 11(4):454–8. <https://doi.org/10.1177/1747493016643851> PMID: 27073187
43. Kwakkel G, Lannin NA, Borschmann K, English C, Ali M, Churilov L, et al. Standardized Measurement of Sensorimotor Recovery in Stroke Trials: Consensus-Based Core Recommendations from the Stroke Recovery and Rehabilitation Roundtable. *Neurorehabil Neural Repair*. 2017; 31(9):784–92. <https://doi.org/10.1177/1545968317732662> PMID: 28934918
44. Walker MF, Hoffmann TC, Brady MC, Dean CM, Eng JJ, Farrin AJ, et al. Improving the Development, Monitoring and Reporting of Stroke Rehabilitation Research: Consensus-Based Core Recommendations from the Stroke Recovery and Rehabilitation Roundtable\*. *Neurorehabil Neural Repair*. 2017; 11(5):472–9. <https://doi.org/https%3A//doi.org/10.1177/1545968317732686> PMID: 29233072
45. Ivers NM, Halperin IJ, Barnsley J, Grimshaw JM, Shah BR, Tu K, et al. Allocation techniques for balance at baseline in cluster randomized trials: a methodological review. *Trials*. 2012; 13(120):1–9. <https://doi.org/10.1186/1745-6215-13-120> PMID: 22853820