

DEVELOPMENT AND VALIDITY OF AN INNOVATIVE TEST TO ASSESS GUIDELINE-CONSISTENT CLINICAL REASONING BY PHYSICAL THERAPISTS IN STROKE REHABILITATION

Nicoline M. OTTERMAN, PhD^{1,2}, Marjo MAAS, PhD^{3,4}, Sven K. SCHIEMANCK, MD, PhD⁵, Philip J. VAN DER WEES, PhD⁴ and Gert KWAKKEL, PhD^{6,7,8}

From the ¹Department of Rehabilitation, Academic Medical Center, University of Amsterdam, ²Basalt Rehabilitation Centre The Hague, ³University of Applied Sciences, HAN Arnhem, ⁴Radboud University Medical Center, Radboud Institute for Health Sciences, Scientific Center for Quality of Healthcare, Nijmegen, ⁵Department of Rehabilitation, Leiden University Medical Center, ⁶Department of Neurorehabilitation, Reade Centre for Rehabilitation and Rheumatology, ⁷Department of Rehabilitation Medicine, VU University Medical Center, MOVE Research Institute, Amsterdam, The Netherlands and ⁸Department of Physical Therapy and Human Movement Sciences, Northwestern University, Chicago, IL, USA

Objective: To evaluate the validity of a script concordance test to assess guideline-consistent clinical reasoning by physical therapists in stroke rehabilitation, and to identify critical features of physical therapists specializing in stroke rehabilitation.

Methods: A script concordance test was developed according to current standards. Four subgroups of physical therapists (those specializing in neurology, those focusing on neurology or geriatrics, other, and non-specialized undergraduate students) were asked to complete the test. The construct validity of the script concordance test was evaluated with 1-way analysis of variance (ANOVA) to estimate differences between subgroups. Associations between physical therapist characteristics, and script concordance test scores were analysed with bivariate regression analysis followed by multivariate analyses.

Results: The script concordance test, with 59 items, was completed by 211 physical therapists. ANOVA analysis showed statistically significant differences between the script concordance test scores of the 4 groups ($p < 0.001$), with higher scores by the physical therapists specializing in neurology compared with the other, non-specialized, subgroups. The multivariate analysis showed that better guideline knowledge ($B = 1.07$; $CI = 0.48-1.65$; $p = < 0.001$), successful completion of the Dutch Neurorehabilitation course ($B = 4.1$; $CI = 1.37-6.87$; $p = 0.003$), and participation in professional development activities ($B = 2.4$; $CI = 0.05-4.68$; $p = 0.046$) were associated with higher script concordance test scores.

Conclusion: The script concordance test has good construct validity. Greater self-reported guideline knowledge, successful completion of the post-bachelor Dutch Neurorehabilitation course, as well as systematic participation in professional development activities facilitate important factors that enhance specialization. The script concordance test is a valid feedback tool for physical therapists to support professional development in the domain of stroke rehabilitation.

Key words: stroke; clinical competence; script concordance test; physical therapy.

Accepted May 3, 2019; Epub ahead of print May 17, 2019

J Rehabil Med 2019; 51: 418–425

LAY ABSTRACT

Adherence of physical therapists to scientific recommendations can improve care for stroke patients. This study developed a test to measure the use of scientific recommendations in the clinical care that physical therapists provide to patients with stroke. The study also identified critical features of therapists who specialize in stroke rehabilitation. It was concluded that the test could indeed distinguish therapists specializing in stroke rehabilitation. The study also found that higher knowledge levels, successful completion of the Dutch Neurorehabilitation course and systematic participation in professional development activities, such as multidisciplinary team meetings and regional case meetings, stimulate the use of scientific recommendations and thus enhance specialization. The test developed in this study is a potential tool to support improvement in continued education to promote professional development of physical therapists in stroke rehabilitation.

Correspondence address: Gert Kwakkel, VU University Medical Center Department of Rehabilitation, De Boelelaan 1117, 1081 HV Amsterdam The Netherlands. E-mail: g.kwakkel@vumc.nl

Several systematic reviews have shown that the evidence and underlying key assumptions for applying neurological treatment approaches, such as Bobath, are weak (1–3). Up to 2008, several professional debates were held in the Netherlands to encourage a more eclectic, evidence-based treatment approach, using the International Classification of Functioning (ICF) as a framework and based on the current understanding of mechanisms that drive stroke recovery and functional prognosis. This eclectic approach allows us to apply new interventions (3) and innovative treatment strategies, based on recently developed theories about sensorimotor recovery for which evidence has been found in the field of stroke rehabilitation (4, 5). The growing and ever-changing amount of evidence in the field of stroke rehabilitation covers approximately 53 different interventions (2), indicating that stroke rehabilitation is a continuum of complexity, with highly complex and specific post-stroke treatments (1). This complexity challenges the use of evidence in the clinical reasoning process of physical therapists (PTs). We

are of the opinion that adherence of PTs to changing clinical practice guidelines in stroke rehabilitation (6, 7), including guideline-consistent clinical reasoning, requires special skills that need to be trained in postgraduate specialization in stroke rehabilitation in order to improve care for stroke patients. Supported by a grant from the Dutch National Institute of Health, a 1-year course, called the Dutch Post-Stroke Neurorehabilitation Course, was started in the Netherlands (3).

Clinical reasoning is a vital component of clinical competence (8), and its importance is recognized in major policy documents about professional development. Higgs and co-workers (9), as well as Edwards and colleagues (10), defined clinical reasoning as a process in which the therapist, interacting with the patient and others, such as family members or other healthcare providers, helps to structure meaning, goals, and health management strategies based on clinical data, patient wishes and choices, and professional judgement and knowledge (9, 10). Systematic application of scientific evidence in the clinical reasoning process is thought to improve the effectiveness and quality of physical therapy practice. We are therefore of the opinion that assessment of the consistent use of clinical practice guidelines (CPGs) in clinical reasoning can provide feedback on the use of guidelines and can identify areas of improvement (11). However, valid measures of guideline-consistent clinical reasoning are scarce in postgraduate physical therapy practice (Box 1).

The script concordance test (SCT) is a written test, based on brief clinical scenarios, to assess clinical reasoning in a context of uncertainty. This test measures the extent to which the clinical reasoning of an individual professional matches that of a group of experts (8, 12, 13). Several studies on the assessment of clinical reasoning in the medical domain have shown that the SCT is a valid assessment tool across a variety of continuing professional development activities (12,

14). However, no research evidence is available on the validity of the SCT as a tool to assess the clinical reasoning of PTs. We are of the opinion that this tool might help us examine the use of guidelines in the clinical reasoning process of PTs in stroke rehabilitation.

This paper first describes the development of an SCT for physical therapy, focusing on the diagnosis, clinical assessment, neurological and functional prognosis for outcome, and treatment in accordance with the Dutch CPG on Stroke in physical therapy (further referred to as CPG Stroke) (2, 15). The study then evaluates the validity of the SCT, defined as the degree to which the scores (17) of the SCT are consistent with our hypothesis that there are differences between the following 4 groups: (i) PTs specializing in neurology; (ii) PTs focusing on neurology or geriatrics; (iii) other or non-specialized PTs; (iv) undergraduate physical therapy students. It was hypothesized that PTs specializing in neurology could be distinguished by their higher SCT score. Furthermore, it was identified whether knowledge, expertise and professional development of PTs were related to guideline-consistent clinical reasoning, in order to identify the critical features of PTs who specialize in stroke rehabilitation.

METHODS

Development of the script concordance test

Development of the content. The SCT was developed according to the guidelines published by the Association for Medical Education in Europe (AMEE) for SCT construction (8, 12, 13) (Fig. 1). First, the project group carefully determined the scope of the SCT, including operationalization of its purpose and the focus of the assessment (Table I). Secondly, a test blueprint was created to bolster the content validity of the SCT. The creation of the test blueprint was a dynamic process, running in parallel with the construction of the test, based on scientific reviews of stroke rehabilitation (1, 2, 15, 16). Thirdly, clinical scenarios were constructed, also known as item vignettes or case vignettes, followed by a set of 2–3 questions, which could be scored on a 5-point Likert scale. The aim was to construct 20–25 vignettes with a total of 60–75 questions in order to achieve sufficient score reliability (12). Two authors (MM and NMO) were responsible for developing draft vignettes. The authenticity of the clinical

Box 1. Glossary

Specialist physical therapist: a physical therapist who has formally demonstrated an ability to apply advanced clinical competence in a defined clinical area, within the scope of practice recognized as physical therapy. A specialist physical therapist will work primarily in a specific area of clinical and/or teaching practice, but would be expected to also be involved in research and evaluation and practice/service development relevant to their practice setting. (10).

Clinical competence: the ability to perform a specific task in a manner that yields desirable outcomes in healthcare. This implies the ability to apply knowledge, skills and abilities successfully to new situations, as well as to familiar tasks for which prescribed standards exist. In addition to skills, clinical reasoning is a vital component of clinical competence.

Guideline – consistent clinical reasoning: the consistent use of recommendations from (inter)national guidelines on thinking and decision-making in professional practice to guide practice activities, such as establishing a diagnosis, and the choice of measurement instruments and interventions. (9).

Test blueprint: a clear framework with specification of the test. It contains the topics to be included in the test, as well as cognitive dimensions and type of questions. The purpose of a test blueprint is to achieve content validity.

Table I. General principles that defined the scope of the script concordance test and on which the assessment is based

Construct to be assessed	Guideline-consistent clinical reasoning as an aspect of specialization
Assessment method	Script concordance test
Purpose of test	Reflecting on guideline-consistent clinical reasoning that can enhance learning. Discriminating between professionals whose reasoning matches that of experts and those whose reasoning does not.
Target group	Practicing physical therapists (postgraduate)
Knowledge domain	Stroke rehabilitation
Focus of the test	Diagnostics and functional prognosis supported by measurement instruments. Treatment planning; goal-setting and choice of intervention

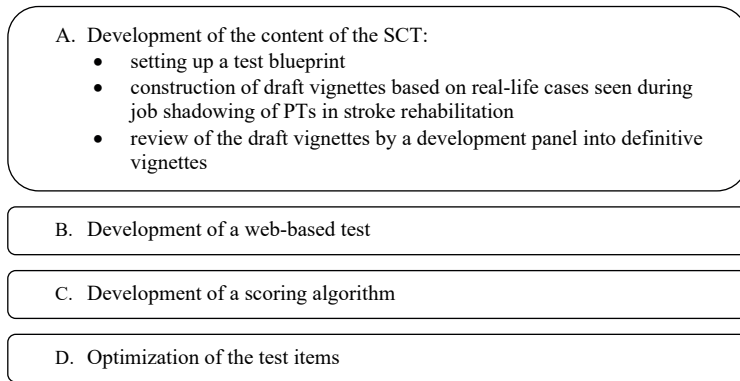


Fig. 1. Development of the script concordance test (SCT) in 4 phases. Flowchart showing the phases in the development of the SCT. PTs: physical therapists.

scenarios was enhanced by studying real-life case examples, by job shadowing the first author in different work settings across the continuum of stroke care. The draft vignettes were reviewed by a development panel, consisting of 4 PTs who had participated in the development of the CPG Stroke. Consensus on relevance, clarity and content of the “cases” was reached in 2 e-mail rounds and 2 in-person consensus rounds.

Development of a web-based test. The final version of the case scenarios and items for the SCT were agreed upon by the authors and the development panel, and this version was programmed in a web-based test.

Development of a scoring algorithm. To develop the scoring algorithm, a reference panel consisting of 15 members was invited to complete the test. Members of the project group could nominate PTs from their network for the reference panel. Individuals were selected if they met all of the following 4 criteria: (i) registered in the Central Quality Register for Physical Therapy; (ii) high level of guideline knowledge and use of the CPG Stroke guideline in clinical practice, based on their participation in the development of the CPG Stroke and/or teaching a course in neurorehabilitation for PTs in which the CPG Stroke was used; (iii) consensus of at least 3 members of the project group and 3 members of the development panel about their level of expertise; and (iv) providing informed consent for participation in the study. Scores for each question were computed from the answers chosen by the reference panel, as proposed in the AMEE guideline (12). Credit for each answer was transformed proportionally to obtain a maximum score of 1 credit for the modal answer from the reference panel for each item, a score of 0 credits for an answer that was not selected by any of the reference panel and partial credit for an answer other than the modal answer. A web-based calculator developed by the University of Montreal was used to analyse the reference panel’s response and construct the scoring algorithm (16).

Optimization of the test items. In establishing the final version of an SCT, different quality assessment strategies for item optimization have been described. A stepwise quality assessment was performed based on the AMEE guideline (12) on SCT development.

First, the variability of the reference panel responses was assessed. The variability among the members of the reference panel has been shown to be a key determinant of the discriminatory power of an SCT (12). Ideally, SCT questions produce a range of expert responses clustered around a modal answer. Questions with unanimity or with a broad distribution of responses are considered to have low quality. The quality of answers was rated using criteria based on the AMEE guideline (12), as follows: (i) high

quality: a range around a modal answer (maximum 3 different response categories); (ii) good quality: a maximum of 4 different response categories with only one expert choosing the extreme answer; (iii) doubtful quality: the 2 highest scoring answers are more than one response category apart; and (iv) poor quality: broad distribution. The doubtful and poor items were considered to be performing poorly in the SCT.

Secondly, the item-total correlation coefficient of the subject responses was assessed, which provided an estimate of each item’s discriminative capacity. A negative or low item-total correlation contributes minimally or not at all to the reliability of the test, although it can also reflect the heterogeneity of clinical competence or the nature of the domain tested. Therefore, it should be carefully considered whether the items with negative or below 0.05 item-total correlations should be discarded.

Thirdly, the content validity as perceived by the reference panel was assessed, rated on a 5-point Likert-scale ranging from (fully) disagree (0) to (fully) agree (5). For each item, the percentage of the members of the reference panel who judged that this item was an adequate reflection of guideline-consistent clinical reasoning, and was relevant for daily practice for patients with stroke, was calculated. An arbitrary cut-off point of 65% agreement or full agreement on this item was used. Percentage scores below this cut-off point were considered to indicate low content validity.

The items that performed poorly on 2 of the 3 levels of quality assessed were presented to the development panel. If 75% of the panel recommended removal of this item, it was discarded.

The total score of the optimized SCT was the sum of the credits of the remaining items, expressed as a percentage of the maximum score.

Recruitment and study sample

An undirected recruitment campaign, using an e-mail sent in February 2015 to 1,704 potential participants, was performed. Post-graduate PTs ($n=728$) were approached via a Dutch national education institute for allied health professionals (Nederlands Paramedisch Instituut). This was a sample of PTs with a variety of fields of interest, as recorded by the institute, such as sports, musculoskeletal, neurology, cardiology and oncology. Physical therapy students ($n=976$) at 7 universities of applied sciences with a physical therapy programme were also approached. All received a reminder e-mail in March 2015. After a positive response to the recruitment mail, a participant received a log-in code for the web-based SCT. After completing the SCT, participants were assigned to 1 of 4 groups based on specialization. Since there is no formal registry of PTs specializing in neurology in the Netherlands held by an institution or society, the authors defined a classification based on therapist characteristics. The first group consisted of a PTs specializing in neurology who met the specialization criteria stated in the CPG Stroke (i.e. treatment volume of at least 5 unique stroke patients a year, completion of the postgraduate course on stroke rehabilitation, participation in professional development activities in the field of stroke, and self-report of neurology being their main specialization). The second group consisted of PTs with a self-reported focus on geriatrics or neurology who did not meet all criteria for the group specializing in neurology. The third group consisted of PTs with other specializations (e.g. mus-

culoskeletal or cardiovascular), or no specialization. The fourth group consisted of undergraduate physical therapy students.

The SCT was web-based and available for completion by the participants from February until June 2015. In addition, participants completed a survey to identify demographic characteristics and clinical expertise. By participating, PTs could obtain continuing education credits for the Dutch Central Quality Register for Physical Therapy. No incentive was available for the physical therapy students.

Statistical analysis

Data for all participants were entered into a computer database and analysed with the IBM SPSS statistical package (version 21.0). Descriptive statistics were used to present the characteristics of the subjects and the SCT score.

Construct validity. Construct validity was defined as the degree to which the scores (17) of the SCT were consistent with our hypothesis that there are differences between the 4 groups, so that PTs specializing in neurology could be distinguished by their higher SCT score. The construct validity of the SCT was analysed with 1-way ANOVA to estimate differences between subgroups with a Bonferroni post-hoc analysis.

Identification of critical features of PTs specializing in stroke rehabilitation. To identify physical therapy characteristics associated with guideline-consistent clinical reasoning in stroke rehabilitation, the optimized SCT score, expressed as percentages of the total score, was used as the dependent variable. To avoid random allocation of physical therapy characteristics in the regression model, 10 hypotheses were formulated based on the expertise of the project group and development panel. It was hypothesized that the following features of clinical competence would be related to a higher SCT score:

- Higher level of guideline knowledge: defined as higher self-reported clinical practice guideline knowledge, using a numerical rating scale.
- Being acquainted with international clinical practice guideline(s) in stroke rehabilitation. Acquaintance was positive when 1 or more international clinical practice guideline(s) were mentioned in response to the open-ended question.
- More years of employment within specialized physical therapy.
- Delivering care within a team, with teamwork defined as working with at least 3 different healthcare professionals, including structured multidisciplinary team meetings to discuss patients' treatment plans.
- A higher number of unique stroke patients treated in the last 12 months, as reported by the PT.
- A higher frequency of reading scientific literature about stroke care per year, scientific literature being defined as international peer-reviewed journals.
- University-level education, defined as having successfully completed a research Master's degree or a PhD programme.
- Having attended in-depth training on neurorehabilitation, defined as successfully having completed the Dutch Neurorehabilitation course.
- Participation in professional development activities, defined as systematic participation in teaching on neurorehabilitation, formalized community stroke networks, or case discussion meetings or peer feedback groups for neurology of the Royal Dutch Society for Physical Therapy.

Associations between physical therapy characteristics and SCT score were first analysed in a bivariate regression analysis to identify statistically significant independent determinants ($p < 0.1$) of the SCT score. Subsequently, multicollinearity was assessed, and only physical therapy characteristics with a correlation coefficient > 0.7 , VIF > 10 or Tolerance < 0.2 were selected, to prevent over-parametrization of the prediction model. Finally, the remaining characteristics were included in a multivariate, forward-selection linear regression analysis followed by a backwards selection method. Standardized and unstandardized coefficients (B) were estimated, with 95% confidence intervals (95% CI) for the unstandardized coefficients. Only physical therapy characteristics with a 2-tailed statistically significant level with a p -value < 0.05 in both methods were considered to make a significant contribution to the multivariable regression model.

RESULTS

Development of the content-valid script concordance test

The SCT developed contained 59 items, using a fixed questioning format, nested within 21 clinical scenarios. All criteria stated in the AMEE guideline (12) were met and the items were sufficiently distributed across the different domains of the test blueprint. An example item is presented in Box 2. A reference panel of 15 physical therapy experts on stroke rehabilitation developed the scoring algorithm. After the quality assessment, 12 items were assumed to perform poorly and the development panel selected 9 of these 12 items to be discarded. After this optimization procedure, 50 items in 21 clinical scenarios were used to calculate the total SCT score for each participant, expressed as a percentage of the maximum score.

Box 2. Script concordance test (SCT) example. Example item from the SCT developed for physical therapy in stroke patients, translated from the Dutch. This concerns the use of a measurement instrument. The bold text represents the fixed format of the script concordance question.

Case description: Mr Klaas was hospitalized with an ischaemic stroke in the left hemisphere. There is no indication for further medical treatment. The physician wants to discharge him home, and seeks your advice. Mr Klaas lives with his partner in an apartment, accessible by stairs as well as by elevator.

Clinical decision-making: Your initial advice is to discharge Mr Klaas home, since he shows maximum scores on the measurement instruments recommended in the clinical practice guideline, and he has no signs of higher cortical impairments.

The following new information is available: Just before discharge, Mr Klaas suffers a recurrent stroke with the same localization. The Motricity Index score after this second stroke is 26/25/25 for his arm, and 33/25/25 for his leg. Furthermore, specific problems of balance are seen and the Berg Balance Score is 50. Mr Klaas is able to walk independently and safely at the hospital department.

As a result, your initial advice becomes:

- Unlikely
- Less likely
- No more nor less likely
- More likely
- Very likely

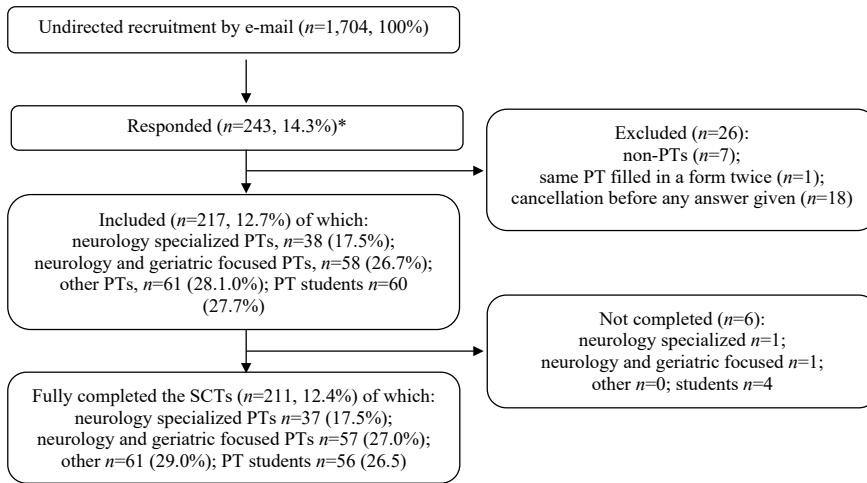


Fig. 2. Recruitment of participants. Flowchart of the physical therapists (PTs) invited and analysed in this study. * $n=11$ missing due to cancelled forms or incorrect log-in before providing information on the group. SCT: script concordance test.

Participants

A total of 211 participants fully completed the SCT, divided into 4 different groups of PTs (Fig. 2). Participants' characteristics, experience, education and SCT scores are summarized in Table II. Complete datasets for the regression analysis were available for 210 subjects.

Construct validity of Script concordance test

ANOVA analysis showed statistically significant differences in SCT scores between the 4 groups ($p < 0.001$). Levene's test revealed that population variances were equal ($p > 0.05$). The Bonferroni test showed statistically significant differences ($p < 0.01$) between the mean SCT score of the group specializing in neurology and the

Table II. Participants' characteristics. Note: there are small variations in the numbers of missing values, with a maximum of 2 subjects missing per group per characteristic

	Groups			
	Neurology specialization	Neurology or geriatric focus	Other or no specialization	Student PT
Participants, n	37	57	61	56
<i>Characteristics</i>				
Sex, %, male	81.1	75.4	73.8	62.5
Age, mean (SD)	42 (10)	40 (11)	41 (13)	21 (2.5)
<i>Experience</i>				
Years of employment as a PT, median (IQR)	16 (19)	13 (18)	16 (24)	-
Years of employment within group specialization, median (IQR)	14 (15)	9 (17)	9 (12)	-
Number of unique stroke patients treated in past 12 months, median (IQR)	40 (9)	20 (25)	2 (9)	-
Current work setting, %				
Community practice	13.5	10.5	49.2	-
Nursing home	27.1	59.6	16.4	-
Rehabilitation Center	21.6	12.3	8.2	-
Hospital	37.8	14.0	26.2	-
Working in a team, %	86.5	91.2	67.2	-
<i>Education</i>				
Entry level education, %				
Not yet graduated in PT	-	-	-	98.2
Bachelor's degree in PT	94.6	93.0	96.7	-
Master's degree or PhD	5.4	7.0	3.3	1.8
Self-reported rating of guideline knowledge (0-10), median (IQR)	7 (1)	7 (1)	6 (1)	6 (1)
Frequency of reading scientific literature on stroke care per year, median (IQR)	5 (8)	2 (5)	0 (2)	0 (1)
Acquainted with international clinical practice guideline(s) in stroke rehabilitation, %	13.5	0	0	0
Dutch Neurorehabilitation course successfully completed, %	64.9	33.3	1.6	-
Participating in additional professional development activities, %	100	47.5	16.4	-
Registered in Dutch Central Quality Register for Physical Therapy, %	94.6	96.5	100	-
<i>Script concordance test (SCT)</i>				
Score in % of the maximum score, means (SD)	69.4 (6.7)	65.2 (7.6)	59.0 (6.3)	58.6 (7.0)
Time spent to fill in the SCTPTS by all groups, min, mean (SD)	140 (70)			

SD: standard deviation; IQR: interquartile range; PT: physical therapists.

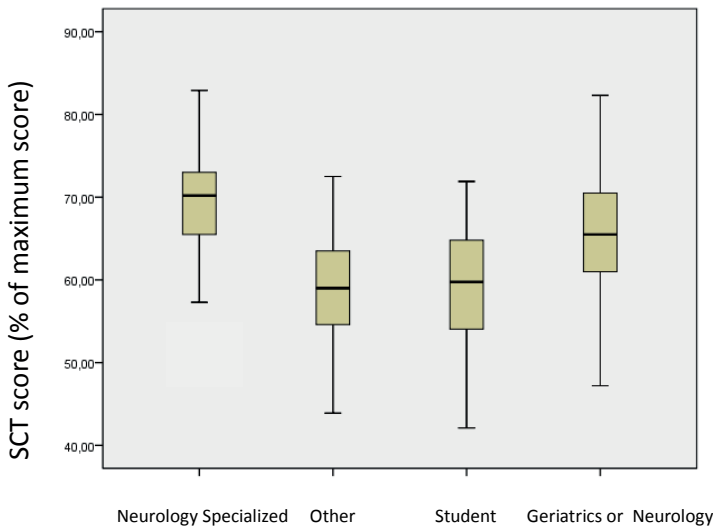


Fig. 3. Boxplot of script concordance test (SCT) score per group, showing the median score (the central horizontal line), the middle 50% of score (the box), and the minimum and maximum scores (bottom and top horizontal lines). Neurology Specialized: PTs specializing in neurology according to criteria of the Dutch Stroke Guideline; Other, other specializations or non-specialized (e.g. musculoskeletal or cardiovascular specialization); Student, students of PT: Geriatrics or Neurology focus; PTs focusing on geriatrics or neurology who do not meet all criteria for the neurology specialized group.

graduated PTs with a different specialization, as well as with the physical therapy students and the group focusing on neurology or geriatrics. The mean differences between the students and the other specialized PTs did not achieve statistical significance ($p=0.3$). This is illustrated in a boxplot of SCT score per group (Fig. 3)

Identification of critical features of physical therapists specializing in stroke rehabilitation

The bivariate analysis showed the associations between the SCT score and six PT characteristics: Self-reported ra-

ting of guideline knowledge, the frequency of reading scientific literature, the number of unique stroke patients treated in the last 12 months, acquaintance with international clinical guidelines, completion of the Dutch Neurorehabilitation course, and participation in professional development activities ($p<0.1$) (Table III). Since no multicollinearity was detected between the determinants, all determinants were included in the multivariate analysis

The multivariate analysis, using both forward and backward linear regression, showed that greater guideline knowledge, successful completion of the Dutch Neurorehabilitation course, and participation in professional development activities were statistically significantly associated with a higher SCT score (Table III). In this model, 22.5% (R^2) of the variability of the SCT was accounted for by these 3 factors.

DISCUSSION

To our knowledge, this is the first study in which a SCT was developed for PTs working in stroke rehabilitation to discriminate between levels of specialization in terms of implementing consistent clinical reasoning. The SCT showed good construct validity, as it was able to distinguish PTs specializing in stroke rehabilitation on the basis of their SCT scores. The SCT score reflects the degree to which their use of the Stroke CPG matched that of experts. Furthermore, our data show that greater perceived knowledge of the stroke guidelines, successful completion of the Dutch Neurorehabilitation course, and systematic participa-

Table III. Bivariate and multivariate linear regression analysis

	Bivariate analysis					Multivariate analysis ($n=210$)				
	Standardized	Unstandardized				Standardized	Unstandardized			
	Beta	B	SE	p -value	95% CI	Beta	B	SE	p -value	95% CI
Self-reported rating of guideline knowledge*	0.41	1.69	0.26	0.000	1.17–2.21	0.26	1.07	0.30	<0.001	0.48–1.65
Frequency of reading scientific literature on stroke care per year*	0.21	0.23	0.08	0.002	0.08–0.38					
Years of employment within a specialization	0.08	0.07	0.06	0.247	-0.05–0.19					
Number of unique stroke patients treated per year*	0.20	0.02	0.01	0.003	0.01–0.04					
Acquainted with international guideline (s) in stroke rehabilitation*	0.20	10.61	3.59	0.003	3.54–17.68					
Working in a team ($n=133$)	-0.02	-0.001	0.003	0.760	-0.01–0.01					
University education ($n=9$)	0.10	3.78	2.74	0.170	-1.63–9.18					
Dutch Neurorehabilitation course successfully completed* ($n=44$)	0.38	7.42	1.27	0.000	4.92–9.93	0.21	4.1	1.39	0.003	1.37–6.87
Participating in additional professional development activities* ($n=78$)	0.34	5.62	1.09	0.000	3.48–7.76	0.14	2.4	1.17	0.046	0.05–4.68
Constant							54.62	1.60	<0.001	51.47–57.77

*Significant determinants of the score on the script concordance test for physical therapists in stroke care ($p<0.1$). n : number of participants; B: unstandardized beta coefficient; SE: standard error of the estimate; 95% CI: 95% confidence interval.

tion in professional development activities (i.e. knowledge about evidence-based guidelines, completion of the post-Bachelor's Dutch Neurorehabilitation course, and participation in professional activities, such as teaching, participation in networks, case discussion meetings) did explain a substantial part of the variance of outcome. In other words, each of these 3 factors contributed significantly to higher SCT scores. This finding suggests that education and active participation in knowledge transfer are important aspects that drive specialization experts in stroke rehabilitation.

A strength of this study was the thorough and careful development of the case vignettes for the SCT test. All steps were discussed with professionals in the field of stroke rehabilitation and experts in educational sciences. In addition, the web-based design makes it easy to access and use. Another strength of this study is the measurement of clinical reasoning; the development of a valid assessment tool contributes to the body of knowledge on clinical reasoning and to the improvement of clinical reasoning, which is a vital component of clinical competence.

Limitations

The current study had some limitations. First, the sample size was relatively small, mainly restricted by the limited number of participants in the group specializing in neurology. Secondly, some PTs reported that items were multi-interpretable. The literature shows that clinical reasoning strategies among healthcare professionals differ, depending on the content and the context of the clinical problem. We assume that the empirical-analytical approach of hypothetico-deductive reasoning does not always match the reasoning strategies of participants. Thirdly, the SCT we have developed is relevant only for a limited period, since evidence and healthcare evolve. Therefore, an SCT should be a "living test", which is adjusted to fit new evidence that becomes available and new developments in healthcare, such as precision medicine. The SCT offers the opportunity to take variability into account; for instance, by presenting the same reasoning dilemmas in different individual contexts of clinical uncertainty. However, it should be investigated whether the individual variability that precision medicine accounts for is also assessed validly with the SCT. Fourthly, the SCT could have resulted in greater contrast if the variability among the answers of the reference panel could be reduced, since it seems that there was little consensus between panel members on some items. A follow-up consensus round in the reference panel to construct the scoring algorithm might reveal if outliers can be reduced and the validity increased further. We

recommend this as an extra step in constructing an SCT. Furthermore, an SCT is not useful for all guidelines. The variability in the answers relates to the level of evidence, therefore guidelines that lack a high level of evidence are prone to variability in the scoring by experts and thus in the scoring algorithm. This will translate into low contrasts in SCT scores between PTs, making the SCT less valid for discrimination of experts from non-experts in the field.

Other validation studies of SCTs in the medical domain have examined the differences between students, residents and specialists (18–21). The total number of questions, and thus the time needed for completion, varied from 40 questions in 30 min to 153 questions in 3 h. These mean scores (50–80) from these studies found similar validity and is comparable to the current study. The authors of these studies concluded that the SCT is an innovative test that can distinguish between different levels of clinical reasoning, and that an SCT is a potential tool for professional development. The current study confirms that SCT is also applicable in the field of physical therapy.

These findings suggest that completion of the Dutch Neurorehabilitation course is associated with higher SCT scores and thus stimulates consistent integration of the CPG Stroke in clinical reasoning. Explicit learning methods, such as peer assessment, facilitate the integration of knowledge in clinical reasoning. This is crucial to learning how to integrate knowledge in clinical reasoning, since this is the way to cope with new rapidly growing evidence.

In contrast to physical therapy specializations, such as manual therapy and paediatric physical therapy, the Royal Dutch Society of Physical Therapy does not keep a country-wide registry of PTs working in the field of neurorehabilitation in the Netherlands. To support continued development, we suggest developing a postgraduate specialization programme for PTs specializing in neurology in the Netherlands. The programme could include active learning methods and systematic continuous professional development activities with built-in feedback and peer assessment in order to provide a full continuous quality improvement cycle. Highly motivated PTs can join the programme and make their continuous education activities transparent to different stakeholders. We believe that the construction of such a programme should be initiated by the physical therapy profession, with the aim of creating a learning environment.

Future research should focus on improving the psychometric properties of the SCT and its clinical use. It would be of interest to study whether the test is responsive to individual scores and subscores. This could reveal the components of clinical reasoning and

the area of knowledge that an individual PT excels in, or identify gaps in knowledge that should be improved. This would allow PTs to give direction to and individualize their development. Ideally, the feedback should create intrinsic motivation to improve and excel. Of special interest in the future, will be the study of the relationship between the clinical reasoning of PTs and clinical and patient-reported outcomes.

Conclusion

We conclude that the SCT is a valid instrument for physical therapy practice, offering the potential to support an increase in guideline-consistent clinical reasoning by PTs in the domain of stroke rehabilitation, in undergraduate education, as well as postgraduate professional development.

ACKNOWLEDGEMENTS

The authors would like to thank the development panel for their contributions to this study: Barbara Harmeling-van der Wel (Erasmus MC, Rotterdam), Jip Kamphuis MSc (ViaReva, Apeldoorn), Roland van Peppen (Hogeschool Utrecht) and Janne Veerbeek (VU Medisch Centrum). For the recruitment of participants to this study we would like to thank Jacqueline Outermans (Hogeschool Utrecht), Anne Griet Brader (Hanze hogeschool Groningen), Corjan Hagendijk (Hogeschool van Rotterdam), Susy Braun (Hogeschool Zuyd), Tjarco Koppelaar (Avans Hogeschool), Miriam Wijbenga and Jaap van den Berg (Hogeschool van Amsterdam) and Marleen Buruma (Nederlands Paramedisch Instituut). Finally, we would like to thank all participating physical therapists and students.

The authors have no conflicts of interests to declare.

REFERENCES

- Langhorne P, Bernhardt J, Kwakkel G. Stroke rehabilitation. *Lancet* 2011; 377: 1693–1702.
- Veerbeek JM, van Wegen E, van Peppen R, van der Wees PJ, Hendriks E, Rietberg M, et al. What is the evidence for physical therapy poststroke? A systematic review and meta-analysis *PLoS One* 2014; 9: e87987.
- Kollen BJ, Lennon S, Lyons B, Wheatley-Smith L, Schepers M, Buurke JH, Halfens J, Geurts AC, Kwakkel G. The effectiveness of the Bobath concept in stroke rehabilitation: what is the evidence? *Stroke* 2009; 40: e89–e97.
- Buma F, Kwakkel G, Ramsey N. Understanding upper limb recovery after stroke. *Restor Neurol Neurosci* 2013; 31: 707–722.
- Kitago T, Krakauer JW. Motor learning principles for neurorehabilitation. *Handb Clin Neurol* 2013; 110: 93–103.
- Van Peppen RP, Maissan FJ, Van Genderen FR, Van Dolder R, Van Meeteren NL. Outcome measures in physiotherapy management of patients with stroke: a survey into self-reported use, and barriers to and facilitators for use. *Physiother Res Int* 2008; 13: 255–270.
- Otterman NM, van der Wees PJ, Bernhardt J, Kwakkel G. Physical therapists' guideline adherence on early mobilization and intensity of practice at Dutch acute stroke units: a country-wide survey. *Stroke* 2012; 43: 2395–2401.
- Dory V, Gagnon R, Vanpee D, Charlin B. How to construct and implement script concordance tests: insights from a systematic review. *Med Educ* 2012; 46: 552–563.
- Higgs J, Burn A, Jones M. Integrating clinical reasoning and evidence-based practice. *AACN Clin Issues* 2001; 12: 482–490.
- Edwards I, Jones M, Carr J, Braunack-Mayer A, Jensen GM. Clinical reasoning strategies in physical therapy. *Phys Ther* 2004; 84: 312–30; discussion 331–335.
- Harrison CJ, Konings KD, Schuwirth L, Wass V, van der Vleuten C. Barriers to the uptake and use of feedback in the context of summative assessment. *Adv Health Sci Educ Theory Pract* 2015; 20: 229–245.
- Lubarsky S, Dory V, Duggan P, Gagnon R, Charlin B. Script concordance testing: from theory to practice: AMEE guide no. 75. *Med Teach* 2013; 35: 184–193.
- Fournier JP, Demeester A, Charlin B. Script concordance tests: guidelines for construction. *BMC Med Inform Decis Mak* 2008; 8: 18.
- Lubarsky S, Chalk C, Kazitani D, Gagnon R, Charlin B. The script concordance test: a new tool assessing clinical judgement in neurology. *Can J Neurol Sci* 2009; 36: 326–331.
- Royal Dutch Society of Physical Therapists, Dutch Clinical Practice Guideline Stroke from the Royal Dutch Society of Physical Therapists. 2014 [accessed 2015 Feb]. Available from: http://www.fysionet-evidencebased.nl/images/pdfs/guidelines_in_english/stroke_practice_guidelines_2014.pdf.
- Université de Montreal, Script Concordance Test Scores [accessed 2015 May 25]. Available from: <http://www.cpass.umontreal.ca/recherche-et-developpement/script-concordance-tests-scts/excel-corrector-program.html>.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63: 737–745.
- Ducos G, Lejus C, Sztark F, Nathan N, Fourcade O, Tack I, et al. The script concordance test in anesthesiology: validation of a new tool for assessing clinical reasoning. *Anaesth Crit Care Pain Med* 2015; 34: 11–15.
- Humbert AJ, Besinger B, Miech EJ. Assessing clinical reasoning skills in scenarios of uncertainty: convergent validity for a script concordance test in an emergency medicine clerkship and residency. *Acad Emerg Med* 2011; 18: 627–634.
- Lambert C, Gagnon R, Nguyen D, Charlin B. The script concordance test in radiation oncology: validation study of a new tool to assess clinical reasoning. *Radiat Oncol* 2009; 4: 7–717X-4-7.
- Nouh T, Boutros M, Gagnon R, Reid S, Leslie K, Pace D, et al. The script concordance test as a measure of clinical reasoning: a national validation study. *Am J Surg* 2012; 203: 530–534.