Original Article

Measuring change in arm hand skilled performance in persons with a cervical spinal cord injury: responsiveness of the Van Lieshout Test

AIF Spooren^{1,2,*}, YJM Janssen-Potten¹, MWM Post³, E Kerckhofs⁴, A Nene⁵ and HAM Seelen^{1,6}

¹Institute for Rehabilitation Research, iRv, Hoensbroek, The Netherlands; ²Department of Health Sciences, Provinciale Hogeschool Limburg, Hasselt, Belgium; ³Rehabilitation Centre de Hoogstraat, Utrecht, The Netherlands; ⁴Department of Rehabilitation Science, Vrije Universiteit Brussel, Brussels GQ, Belgium; ⁵Rehabilitation Centre Roessingh, Enschede, The Netherlands; ⁶Rehabilitation Foundation Limburg (SRL), Hoensbroek, The Netherlands

Objectives: To assess whether the Van Lieshout Test (VLT) is responsive to measure changes in arm hand skilled performance (AHSP) over time during active rehabilitation in persons with cervical spinal cord injury (C-SCI) and in different subgroups of persons with C-SCI according to lesion completeness and to lesion level.

Study design: Longitudinal cohort study.

Setting: Spinal Cord Injury Units in eight rehabilitation centres across the Netherlands.

Methods: In the present study, data from a national SCI cohort study are used. Data from the VLT, the Grasp Release Test (GRT), the Functional Independence Measure (FIM) and the Quadriplegia Index of Function (QIF) were recorded at three moments during active rehabilitation: at the start of active rehabilitation (*t*₁), 3 months after *t*₁ (*t*₂) and at discharge (*t*₃). Data have been analysed using three responsiveness measures, that is, the standardized response mean, the effect size and comparison of test scores measured at different times.

Results: The VLT is responsive for the intervals *t*₁–3, *t*₁–2 and *t*₂–3. The VLT can be used to measure changes in AHSP in persons with a C-SCI with an American Spinal Injury Association Impairment Scale (AIS) A or B score or an AIS C or D score, as well as in persons with a C3–C6 lesion or a C7–T1 lesion. The responsiveness of the VLT is significantly correlated to the GRT, but not to the FIM and the QIF.

Conclusion: The VLT is responsive in measuring changes in AHSP during rehabilitation in persons with C-SCI.

Spinal Cord (2006) 44, 772–779. doi:10.1038/sj.sc.3101957; published online 4 July 2006

Keywords: tetraplegia; VLT; responsiveness; arm hand skilled performance

Introduction

Persons with a cervical spinal cord injury (C-SCI) have limited arm hand function causing severe difficulties or even an inability to perform daily skills. The extent of the impaired arm hand function strongly depends on the lesion level and the degree of incompleteness of the lesion. Functional deficits can be measured at the level of (1) impairments of structure and function, (2) activity limitations and (3) participation restrictions as defined in the International Classification of Functioning, Disability and Health (ICF).¹ The level of activities in the ICF includes a broad range of activities from basic activities like grasping and moving objects to complex activities like dressing and grooming. In the present study, the term ‘arm hand function’ (AHF) will refer

to the ICF ‘function’ level, whereas ‘arm hand skilled performance’ (AHSP) or ‘upper extremity skilled performance’ will refer to the level of ‘disability’ or ‘activity’ in accordance with the ICF nomenclature.

A small gain in AHF and AHSP may lead to large differences in self-management and quality of life. Previous research showed that patients with a C-SCI expect a better quality of life if their AHF and AHSP would improve and that improvement in arm and hand is a top priority in comparison to other SCI-associated problems.^{2–5} Therefore, restoring the highest possible level of AHSP and AHF is a very important goal in the rehabilitation of persons with C-SCI.

To establish a good rehabilitation policy with respect to the arm and hand, it is necessary to be able to evaluate and to objectify and quantify the amount of progress in arm and hand, both from a patient’s

*Correspondence: AIF Spooren, Guffenslaan 39, 350U Hasselt, Belgium

motivational point of view and as to provision of information to the patient.

There are some instruments that measure AHF at the ICF function level, like the Manual Muscle Testing⁶ and hand-held dynamometry.⁷ However, for patients these measures remain very abstract as they do not provide information about the actual skills patients will be able to perform. Even for therapists, measurements at function level are not a priority in evaluating arm and hand, as there is a growing interest in measurements at ICF activity level (AHSP) and especially in gauging the quality of movement to assess which strategies patients use to perform a skill.

Despite the vast clinical experience, the objective assessment of upper extremity skilled performance and especially the objective assessment of movement quality is still very limited. Measurements at activity level and more specifically at the level of complex activities encompass, among others, the Functional Independence Measure (FIM)^{8,9} and the Quadriplegia Index of Function (QIF).¹⁰ The FIM is extensively used to evaluate persons with SCI, but it is not specific for C-SCI.¹¹ The QIF has been designed for C-SCI, but both the FIM and the QIF measure more the overall level of independency, which involves complex activities like washing and grooming. They do not provide specific information about AHSP.

The Grasp Release Test (GRT)¹² is a specific hand function test at the ICF level of basic activities and has originally been designed to assess the use of a hand neuroprosthesis in C5–C6 tetraplegic persons, but may be used for all tetraplegic patients. It is a quantitative test that counts the total number of certain objects people can grasp, move and release. Previous research showed a good test–retest reliability of the GRT and the ability of the GRT to detect changes in hand function following tendon transfer and functional electrostimulation.¹³

Because none of the above mentioned-instruments solved the problem of objectively quantifying the quality of the movement of the actual upper extremity performance during different daily skills in persons with a C-SCI, the Van Lieshout Test (VLT) for arm hand performance was developed.¹⁴ The VLT consists of 19 tasks that cover the majority of arm hand functions associated with activity of daily living (ADL) and best represents the level of basic activities according to the ICF. Five basic arm and hand modalities can be distinguished in the VLT, that is, (1) general performance of the arms aimed at movement of the body, (2) positioning and stabilizing the arms, (3) current status regarding opening and closing of the functional hand, (4) grasp and release and (5) manipulation using both thumb and fingers. The execution of each task is scored on a six-point scale (0–5), indicating the quality of performance. To look at the grip strength of the thumb, for example, the way a person can pour a can of water using his thumb is observed. A description of the VLT is given by Post *et al.*¹⁵ and at <http://www.irv.nl/vlt>. Next to the full version of the VLT, which may

be used in comprehensive clinical evaluation, a short 10-item version (VLT-SF) has been designed for research purposes. The VLT-SF highly correlates with the long version of the VLT ($r = 0.90–0.93$)¹⁵ and is used in the present study. Besides objectifying the quality of the movement, the VLT is a starting point for hand function therapy and a tool to assist in clinical decision making as to treatment policy.

The VLT is a new instrument that provides information about strategies of arm and hand patients use to perform an activity. The criterion validity (Spearman's $r = 0.87–0.90$) and the inter-rater reliability (ICC = 0.98–0.99) of the VLT-SF were found to be very good and the internal consistency reliability (Cronbach's $\alpha = 0.88–0.94$) of the VLT-SF was also found to be good.¹⁵ Because of the growing interest of physicians and therapists in measuring progress in clinical status of patients, recent research has put a greater emphasis on a test's capacity to detect change.¹⁶ The present study will look whether the VLT is able to detect changes in AHSP. In general, the psychometric property associated with the ability to measure change is called 'responsiveness' or 'sensitivity to change', although many variations on this definition exist.¹⁷ In the present study, we use the term 'responsiveness' as this term is used more often in literature, although the term 'sensitivity to change' may be more comprehensible in clinical use.

The study of responsiveness is relatively new to the health sciences literature and the discussion as to which analytic strategies should be used is still ongoing. For the 'criteria-based responsiveness', a standard or criterion is used in judging whether there is clinically meaningful change. Many studies use a stable reference group to compare the changing group with.¹⁸ As there was no external criterion or stable group available in the present study, the 'distribution-based responsiveness' will be used.^{18,19}

As clinicians are interested to know whether the VLT can be used to measure changes in AHSP in persons with C-SCI, irrespective of their lesion level and lesion completeness, the present study will describe the responsiveness of the VLT in different subgroups of persons with C-SCI.

At present, the FIM, the QIF and the GRT are the test batteries most often used in persons with C-SCI. Therefore, the responsiveness of the VLT will be investigated alongside the responsiveness of these well-known test batteries as a matter of reference and the responsiveness of the different test batteries will be correlated to each other. However, the aim of the present study is not to compare the responsiveness in terms of 'better' or 'worse' as all test batteries provide different information and therefore comparison may be impeded.

The main hypothesis of the current study is that the VLT may be used to assess changes over time regarding AHSP in persons with C-SCI. The aim of the present study is to answer the following questions: (1) Is the VLT responsive as to AHSP during rehabilitation in persons with cervical SCI? (2) Is the VLT responsive in

persons with a C-SCI with either an A or B score or a C or D score according to the ASIA Impairment scale? (3) Is the VLT responsive in persons with a C3–C6 lesion and persons with a C7–T1 lesion? (4) How does the responsiveness of the VLT relate to the responsiveness of the GRT, the FIM and the QIF in persons with C-SCI?

Methods

Subjects

This study is part of the Dutch research program Physical Strain, Work Capacity and Mechanisms of Restoration of Mobility in the Rehabilitation of Persons with Spinal Cord Injuries (www.fbw.vu.nl/onderzoek/A4zon/ZONenglish). Eight Dutch rehabilitation centres specialized in the rehabilitation of persons with SCI participated in this programme. Patients with an acute SCI have been followed from the onset of rehabilitation to discharge, resulting in three measurement moments as described in the procedure section. The duration of patient influx was 3 years. From the 83 persons with C-SCI who participated in this longitudinal cohort study, 60 persons have been selected. Inclusion criteria consisted of cervical (in)complete SCI (including the T1 level), and age between 18 and 65 years. Exclusion criteria were cervical SCI with severe additional neurological, orthopaedic or rheumatologic diseases, which interfered with the ADL functioning and upper limb skilled performance. Also excluded were persons who performed the upper extremity tests only once ($n=7$), persons with C-SCI with an unimpaired upper extremity performance (because their 'best hand' was at the body side on which the lesion level was below T1) ($n=4$) or persons with C-SCI who were not able to perform upper extremity measurements at all (these were persons with a complete C-SCI at C3 level, for example) ($n=12$). Persons were classified by lesion characteristics according to the International Standards for Neurological and Functional Classification of Spinal Cord Injury.²⁰ During the upper extremity measurements, the best hand was used. Therefore, subjects were categorized by lesion level of the best hand and by level of SCI completeness, that is, persons with an A or B score and persons with a C or D score, according to the American Spinal Injury Association Impairment Scale (AIS).²⁰

Procedures

In each participating rehabilitation centre, one trained research assistant conducted the measurements according to a standardized protocol described in the research programme. Subjects were assessed at three moments during active rehabilitation: at the start of the functional rehabilitation ($t1$), defined as the moment subjects were just able to sit in their wheelchair for at least 3 consecutive hours; 3 months later ($t2$); and at the time of discharge ($t3$). In those cases where the $t2$ measurement

coincided with the time of discharge ($t3$ measurement), the data of $t2$ and $t3$ were identical.

The protocol of the present study was approved by the Medical Ethics Committee of the Institute for Rehabilitation Research and the Rehabilitation Foundation Limburg. All subjects completed a consent form after they had been given information about the test procedures.

Measurements

Tests used to measure specific AHSP at ICF level of basic activities were the VLT and the GRT; other tests measuring the level of independency at ICF level of complex activities were the FIM and the QIF.

VLT-SF The VLT^{15,21} assesses the actual performance of arm hand skills in persons with cervical SCI and looks at the way skills are performed. In the present study, the short version or research version of the VLT (VLT-SF) was used consisting of 10 items aimed at forward reaching, arm extension against gravity, thumb closure, grip function of the thumb, thumb strength, finger closure, finger strength, pen grip, lighting a match and opening a bottle. Some items involved basic arm skills like forward reaching, other items involved hand and finger skills like thumb closure and finger strength. The items pen grip, lighting a match and opening a bottle involved manipulation of objects.

Each item has been scored between 0 and 5 for the best hand, whereby score 5 represented the highest level of performance and score 0 represented that accomplishment of the task was not possible at all. The total score was the sum of all the different item scores with a maximum of 50.

GRT In the GRT,¹² subjects grasp, move and release six different objects as many times as possible in three 30-s trials for each object. Three objects have to be manipulated with palmar grasp and three objects with lateral grasp. The number of attempts, completions and failures are registered. For the present study, the average of the number of completions of the best hand was noted for each item and the total score was the sum of the six item scores. A close relationship between the total score of the GRT and the total score of the VLT-SF was found by Post *et al.*¹⁵

FIM The FIM assesses the level of independence of the persons.^{8,9} In the present study, the Motor score of the FIM was used consisting of 13 items,²² scored on a seven-point scale, varying from total assistance (1) to complete independence (7) with a maximal score of 91. Earlier research found that the FIM was useful in detecting changes in function over time in persons with tetraplegia.²³

QIF QIF is a specific and sensitive instrument to document the functional improvements achieved during the rehabilitation of tetraplegic patients.²⁴ In the current study, the short-form QIF was used, which has a high correlation with the long full version of the QIF.²⁵ The short-form consists of six ADL activities and each item is scored on a five-point scale (0–4) in order of increasing independence with a maximal total score of 24.

Data analysis

One of the programme partners combined the data into a collective database. The database has been verified by checking data with the original forms and by inquiring the research assistants. All test scores were rescaled to 100 to accommodate mutual comparison between test results of the different tests. To normalize the scores, a set maximum score for each test was needed. As the GRT did not have a set maximum score, 200 was considered as maximum score after testing people with unimpaired arm hand performance.

As the ‘distribution-based responsiveness’ has been used in the present study, responsiveness has been assessed using three methods: (1) comparing test scores measured at different times,^{26,27} (2) the standard response mean^{19,28} and (3) the effect size (ES).^{29,30} As to the first method, a Friedman test was used to detect significant changes in consecutive VLT, GRT, FIM and QIF scores across the three measurements between the onset of rehabilitation and discharge with the level of significance set at $\alpha=0.05$. A Wilcoxon matched pairs signed-rank test was used to identify which intervals were significantly different.^{26,27} For multiple comparisons, the Bonferroni adjustment was applied ($\alpha=0.0167$). It was expected that the tests used would be able to measure changes between $t1$ and $t3$. Because of the importance to be able to measure progress during rehabilitation, the present study will look at intervals within the rehabilitation process. As a result, following intervals have been selected:

- $t1-t3$: From start of rehabilitation to discharge.
- $t1-t2$: From start of rehabilitation to 3 months later.
- $t2-t3$: From 3 months after the start of rehabilitation to discharge.

As to the second and the third method, the standardized response mean (SRM), which is the mean change score divided by the standard deviation (SD) of the change score,^{19,28} and the ES, which is the mean change score divided by the SD of the baseline score,^{29,30} have been used. In the present study, we refer to ‘delta values’ as the mean change scores. The SRM and the ES were calculated for the three intervals mentioned above. For the interpretation of the SRM and the ES, a value of 0.20 was considered as a small degree of responsiveness, a value between 0.50 and 0.80 was considered as a moderate degree and a value of 0.80 or more was considered as a large degree of responsiveness.³¹

To relate the responsiveness of the VLT to the other used test batteries, the Spearman’s rho correlation coefficient was performed on the normalized SRM values with the level of significance set at $\alpha=0.05$.^{19,32}

Results

In the present study, 60 persons from a total group of 83 persons were included.

Group composition and mean time between the measurements are presented in Table 1.

Mean change scores or delta values of the three time intervals for the different test batteries and the different groups were calculated and are shown in Table 2 together with the Friedman test results. Values of the Wilcoxon’s signed-rank test, the SRM and the ES for the three time intervals, the different test batteries and the different groups are presented in Table 3.

In five of the 60 persons, $t2$ measurement coincided with the time of discharge and therefore $t2$ and $t3$ data were identical. Missing data (2%) were randomly scattered across patient characteristics and rehabilitation centres.

VLT There was a significant difference in the VLT scores across the three measurements (Friedman, $P<0.001$). More specifically, there was a significant difference between the intervals $t1-3$, $t1-2$ and $t2-3$ (Wilcoxon; $P<0.001$). The SRM for the VLT (SRM_{VLT}) was large for the intervals $t1-3$ and $t1-2$ and moderate for $t2-3$. The ES for the VLT (ES_{VLT}) indicated a moderate degree of responsiveness for the intervals $t1-2$ and $t1-3$.

A–B and C–D Significant differences in VLT scores across three measurements were found for both groups, that is, in persons with a C-SCI with an AIS A or B score (A–B group, $n=34$) and with an AIS C or D score (C–D group, $n=26$) (Friedman, $P<0.001$). The VLT scores of the A–B group as well as the VLT scores of the

Table 1 Population and time intervals

Population		
Age	Mean: 38.9 years	Range: 18.42–64.5
Gender	46 male/14 female	
Level	C3–C6: $n=42$	C7–T1: $n=18$
AIS	A–B: $n=34$	C–D: $n=26$
Time interval	Mean	Range
Days between $t1$ and $t2$	93	64–178
Days between $t1$ and $t3$	288	76–713

AIS, American Spinal Injury Association Impairment Scale

Table 2 Delta values and Friedman test

Test	Total			A-B			C-D			C3-C6			C7-T1		
	Mean	SD	Med	Mean	SD	Med	Mean	SD	Med	Mean	SD	Med	Mean	SD	Med
VLT3-1	24.14	21.86	20	21.59	17.78	21	27.6	26.47	18	24.39	19.68	22	23.56	26.86	17
VLT2-1	18.3	19.19	12	15	14.49	12	22.8	23.77	12	18.54	18.2	12	17.78	21.84	12
VLT3-2	5.83	8.69	4	6.59	9.6	4	4.8	7.33	4	5.85	8.75	4	5.78	8.81	5
VLT(Friedman)	0.000*			0.000*			0.000*			0.000*			0.000*		
GRT3-1	19.59	18.68	15.3	14.46	16	12	26.85	20.09	23	20.15	18.43	15	18.33	19.7	16.5
GRT2-1	14.05	15.08	9.75	9.65	12.48	7	20.29	16.46	15.25	13.57	15.06	9.25	15.11	15.53	9.75
GRT3-2	5.53	8.03	3.75	4.81	8.47	2.75	6.56	7.42	6.5	6.57	8.05	4.5	3.22	7.7	3.25
GRT(Friedman)	0.000*			0.000*			0.000*			0.000*			0.000*		
QIF3-1	39.51	27.6	37.5	35.17	30.69	29.2	45.19	22.26	47.9	31.84	23.98	31.3	57.41	27.81	66.7
QIF2-1	26.32	23.35	20.83	19.85	22.85	12.5	34.78	21.6	37.5	20.83	20.18	14.58	39.12	25.73	35.42
QIF3-2	13.19	17.98	8.33	15.32	20.84	8.33	10.42	13.24	6.25	11.01	13.77	6.25	18.29	25.01	8.33
QIF(Friedman)	0.000*			0.000*			0.000*			0.000*			0.000*		
FIMmot3-1	27.91	18.94	24.4	21.33	16.42	16.5	36.52	18.82	37.4	24.83	18.34	19.8	35.1	18.83	33
FIMmot2-1	18.22	15.65	13.74	11.44	8.18	9.89	27.09	18.59	23.63	17.01	15.89	11.54	21.06	15.14	18.13
FIMmot3-2	9.69	11.46	7.69	9.89	12.91	7.69	9.42	9.47	7.14	7.82	9.32	7.14	14.04	14.74	8.24
FIM(Friedman)	0.000*			0.000*			0.000*			0.000*			0.000*		

VLT, Van Lieshout Test; GRT, Grasp Release Test; QIF, Quadriplegia Index of Function; FIM, Functional Independence Measure; SD, standard deviation

*Identifies statistically significant difference (alpha = 0.05)

All scores are rescaled to 100

Table 3 Responsiveness measurements

Test	Total			A-B			C-D			C3-C6			C7-T1		
	Wilcoxon	SRM	ES	Wilcoxon	SRM	ES	Wilcoxon	SRM	ES	Wilcoxon	SRM	ES	Wilcoxon	SRM	ES
VLT3-1	0.000*	1.1	0.71	0.000*	1.21	0.74	0.000*	1.04	0.78	0.000*	1.24	0.69	0.000*	0.88	0.81
VLT2-1	0.000*	0.95	0.54	0.000*	1.03	0.52	0.000*	0.96	0.64	0.000*	1.02	0.53	0.000*	0.81	0.61
VLT3-2	0.000*	0.67	0.19	0.000*	0.69	0.25	0.002*	0.65	0.2	0.000*	0.67	0.18	0.014*	0.65	0.29
GRT3-1	0.000*	1.05	0.75	0.000*	0.9	0.66	0.000*	1.34	0.97	0.000*	1.09	0.82	0.001*	0.93	0.62
GRT2-1	0.000*	0.93	0.54	0.000*	0.77	0.44	0.000*	1.23	0.74	0.000*	0.9	0.55	0.000*	0.97	0.51
GRT3-2	0.000*	0.69	0.19	0.001*	0.57	0.2	0.001*	0.88	0.26	0.000*	0.82	0.23	0.056	0.42	0.11
QIF3-1	0.000*	1.43	2.18	0.000*	1.15	2.81	0.000*	2.03	2.04	0.000*	1.33	1.61	0.000*	2.08	3.26
QIF2-1	0.000*	1.13	1.38	0.000*	0.87	1.59	0.000*	1.61	1.57	0.000*	1.03	1.05	0.000*	1.52	2.22
QIF3-2	0.000*	0.74	0.4	0.000*	0.73	0.52	0.001*	0.79	0.35	0.000*	0.8	0.34	0.002*	0.73	0.6
FIMmot3-1	0.000*	1.47	2.08	0.000*	1.23	2.01	0.000*	1.94	2.47	0.000*	1.35	2.12	0.000*	1.86	2.08
FIMmot2-1	0.000*	1.16	1.36	0.000*	1.4	1.08	0.000*	1.46	1.83	0.000*	1.07	1.45	0.000*	1.39	1.25
FIMmot3-2	0.000*	0.85	0.42	0.000*	0.77	0.79	0.000*	0.99	0.37	0.000*	0.84	0.34	0.001*	0.95	0.63

SRM, standardized response mean; ES, effect size; VLT, Van Lieshout Test; GRT, Grasp Release Test; QIF, Quadriplegia Index of Function; FIM, Functional Independence Measure

*Identifies statistically significant difference (alpha = 0.0167) responsiveness ES and SRM values: <0.5: small, **0.5–0.8: moderate**; >**0.8: large** (Cohen, 1977)

C-D group showed significant differences in all time intervals (Wilcoxon, $P < 0.003$). In both groups, the SRM_{VLT} was large for the interval $t1-3$ and $t1-2$ and moderate for the interval $t2-3$. As for the total group, the ES_{VLT} of the A-B group and the C-D group demonstrated a moderate degree of responsiveness for

the interval $t1-3$ and $t1-2$ and a small degree of responsiveness in the interval $t2-3$.

C3-C6 and C7-T1 Persons with a C3-C6 lesion ($n = 42$) as well as persons with a C7-T1 lesion

($n=18$) showed significant differences in VLT scores across the three measurements (Friedman, $P<0.001$). The VLT scores of the C3–C6 group as well as the VLT scores of the C7–T1 group showed significant differences in all time intervals (Wilcoxon, $P<0.015$). The SRM_{VLT} of both groups indicated a large degree of responsiveness for the interval $t1-3$ and $t1-2$, whereby the SRM values for the C3–C6 group (1.24 and 1.02) were somewhat higher than for the C7–T1 group (0.88 and 0.81). For the interval $t2-3$, the SRM_{VLT} of both groups indicated a moderate degree of responsiveness. The ES_{VLT} was for the interval $t1-3$ large for the C7–T1 group and moderate for the C3–C6 group. For both groups, the ES_{VLT} was moderate for the interval $t1-2$ and low for the interval $t2-3$.

VLT related to the GRT, the FIM and the QIF As shown in Table 3, the VLT scores showed significant differences in the same time intervals as the scores of the GRT, the FIM and the QIF (Wilcoxon, $P<0.001$). The SRM values indicated a large degree of responsiveness for all test batteries used in the present study for the intervals $t1-3$ and $t1-2$, although there was a slight distinction in the SRM values (see Table 3). The ES_{FIM} and ES_{QIF} indicated a large degree of responsiveness for the intervals $t1-3$ and $t1-2$, whereas the ES_{VLT} and the ES_{GRT} were moderate for these intervals. For the interval $t1-3$, the correlation between the responsiveness of VLT and the GRT was significant (Spearman's $r=0.531$) ($P<0.01$); the correlation between the responsiveness of the VLT and the FIM and the VLT and the QIF was very low (Spearman's $r=0.075$ and $r=0.194$, respectively). Correlation between the responsiveness of the FIM and QIF was also significant (Spearman's $r=0.714$) ($P<0.01$).

Discussion

Research question 1: Is the VLT responsive as to AHSP during rehabilitation in persons with C-SCI?

According to the procedures used to measure responsiveness, it can be concluded that the VLT is indeed responsive for the interval $t1-3$, but also for the intervals $t1-2$ and $t2-3$.

As to the interval $t2-3$, there was a significant difference between the VLT scores at $t2$ and $t3$ (Wilcoxon, $P<0.001$), but SRM_{VLT} and ES_{VLT} were low to moderate. This can be explained by the fact that in some cases ($n=5$) $t2=t3$. As a consequence, the mean delta value of the total group was reduced for this interval. Although both ES and SRM are responsiveness measures, they do not always indicate the same level of responsiveness of the VLT. The ES_{VLT} indicated a moderate degree of responsiveness where the SRM_{VLT} indicated a large degree of responsiveness. This can be explained by the fact that the SDs of the baseline scores (used as the denominator in the ES) were higher than the SDs of the delta scores (used as the denominator in the SRM).

Research question 2: Is the VLT responsive in persons with a C-SCI with either an AIS A–B score or an AIS C–D score?

From the current study, it can be concluded that the VLT is responsive in both groups in all intervals. This means that the VLT can be used to measure changes in AHSP in both persons with C-SCI with an AIS A or B score and in persons with an AIS C or D score during the total rehabilitation process. As a consequence, clinicians, by using the VLT, may be able to objectify changes in the quality of movement in both groups, regardless of the degree of completeness. It was expected that the degree of responsiveness for the C–D group would be higher than for the A–B group as the former often make more progress.^{33–35} However, this hypothesis was not confirmed by the VLT test results.

Research question 3: Is the VLT responsiveness in persons with a C3–C6 lesion and persons with a C7–T1 lesion?

The VLT is responsive in both groups in all intervals. This means that the VLT is able to measure changes in AHSP and to provide information about the changes in quality of movement during the entire rehabilitation process both in persons with a high C-SCI and in persons with a low C-SCI.

Research question 4: How does the responsiveness of the VLT relate to the responsiveness of the GRT, the FIM and the QIF?

In general, it can be concluded that the VLT has, with a few exceptions, the same degree of responsiveness as the GRT, the FIM and the QIF. However, the SRM and ES values were somewhat higher for the FIM and the QIF. This can be explained by the higher variance of the delta scores and the baseline scores of the VLT and the GRT (as shown in Table 2). But from clinical point of view, there is also an explanation of the higher responsiveness values of the FIM and the QIF. The FIM and the QIF both measure the level of independency, which involves more complex activities than activities at basic level as tested by the GRT and the VLT.¹¹ Therefore, they may detect changes that are not solely influenced by changed AHF and AHSP, for example, changes in transfers (measured by the FIM) cannot be attributed to changes in arm and hand only. Furthermore, in the FIM and the QIF, no record is kept whether items are performed with help of assistive technology or whether tasks are performed bimanually, as their aim is to assess independency. In contrast, the VLT and the GRT mainly focus on the best hand and do not allow compensations or assistive technology. As a consequence, the test results of the VLT and the GRT are mainly attributed to the performance of the tested arm and hand. Therefore, less progress is expected in the VLT and the GRT and the responsiveness values may be lower than for the FIM and the QIF.

This clinical explanation was statistically corroborated by the correlation of the SRM values in the

present study. A significant correlation was found between the test batteries, which measure at the same level of AHSP, that is a significant correlation between the VLT and the GRT (which both measure at basic activity level) and a significant correlation between the FIM and the QIF (which measure at complex activity level). The relation between the VLT and the FIM or the QIF was not significant. This means that changes detected by the VLT do not correspond well with changes measured by the FIM or the QIF and correspond partly with changes measured by the GRT. However, although the correlation between the VLT and the GRT was significant, the information both test batteries provide is totally different. The GRT expresses the quantity of how many times a specific task can be performed and focuses mainly on the hand,¹³ whereas the VLT provides information about the quality of a movement, that is the way certain tasks are performed and how arm and hand are involved. In contrast to the VLT, the GRT was already found significantly sensitive to detect changes in the total number of objects that can be handled after surgery as compared to the number of objects handled before surgery.¹³ In the present study, both the VLT and the GRT were found to be responsive to detect changes in hand performance during rehabilitation in persons with C-SCI. However, as they provide different information and changes measured by the VLT correspond partly with changes detected by the GRT, the changes measured by the VLT may provide important additional information to evaluate and objectify AHSP.

Methodological considerations

Some limitations may be identified in the present study. Sixty persons from a total group of 83 persons were included. Although this is not a very large group, it can be considered to be a good representation of the total C-SCI population as the group of 83 persons was the majority of the C-SCI persons admitted in the rehabilitation centres specialized in SCI for the last 3 years in the Netherlands. Other persons with a C-SCI were excluded from the programme because of different reasons, that is, progressive diseases, psychological problems, language problems or refusing to participate in the program. Another limitation was that only the best hand was measured for the VLT and the GRT. As a consequence, some people did not make much progress, because they had reached almost the maximal score with their best hand and other people were excluded because their 'best hand' was unimpaired. The variance of the VLT scores was somewhat high, which might diminish when assessing a larger sample. However, when the group was split into subgroups to answer questions 2 and 3, we could still establish responsiveness for the VLT, indicating the variance was less of a problem in this study.

When evaluating responsiveness, it would be ideal to have two baseline measurements before follow-up measurements.¹⁹ Because this was too demanding for

the participants at that specific time so early in their rehabilitation programme, only one baseline measurement was taken. Furthermore, one might argue that a learning effect may have influenced the responsiveness, that is, the patient being more familiar with tests the second time. However, there are some arguments against this phenomena occurring. Firstly, patients received information about the measurements and had some training trials before measuring to familiarize themselves with the tests. Secondly, the measured tasks were commonly used activities that were also practised during therapy. Thirdly, the time between subsequent measurements was very long, so that the chance that learning effect would take place was very small.

Guyatt *et al*³⁶ defines responsiveness as the ability to detect minimal clinically important differences. Looking at the values of the SRM and the ES, it can be concluded that the VLT is responsive and therefore able to detect changes, without explicitly quantifying the level of clinical importance of the detected changes. However, Snoek *et al*² stated that the improvement in AHF and AHSP is a major issue in the quality of life of tetraplegic patients and that very small improvement in AHF and AHSP is already clinically relevant. The detected change in AHSP by the VLT may indicate a clinically important change. Future research could statistically verify whether this change detected by the VLT is really clinically important.

Conclusion

The VLT is responsive to measure changes in AHSP during rehabilitation in people with a C-SCI and may provide important additional information about changes in AHSP.

Acknowledgements

This study is part of the ZONMW Rehabilitation Research programme 'Physical strain, work capacity and mechanisms of restoration of mobility in the rehabilitation of individuals with spinal cord injury' and was funded through Grant 1450010. Funds were also provided by the Provincial Hogeschool Limburg, Hasselt (Belgium).

References

- 1 Organisation WH. *ICF: International Classification of Functioning for Disability and Health*. Geneva 2001.
- 2 Snoek GJ, IJerman MJ, Hermens HJ, Maxwell D, Biering-Sorensen F. Survey of the needs of patients with spinal cord injury: impact and priority for improvement in hand function in tetraplegics. *Spinal Cord* 2004; **42**: 526–532.
- 3 Welch RD, Lobley SJ, O'Sullivan SB, Freed MM. Functional independence in quadriplegia: critical levels. *Arch Phys Med Rehab* 1986; **67**: 235–240.
- 4 McDowell CL, Rago TA, Gonzalez SM. Tetraplegia. *Hand Clin* 1989; **5**: 343–348.
- 5 Hanson RW, Franklin MR. Sexual loss in relation to other functional losses for spinal cord injured males. *Arch Phys Med Rehab* 1976; **57**: 291–293.

- 6 Council MR. *Aids to the Investigation of the Peripheral Nervous System*. Her Majesty's Stationery Office: London 1976.
- 7 Schwartz S, Cohen ME, Herbison GJ, Shah A. Relationship between two measures of upper extremity strength: manual muscle test compared to hand-held myometry. *Arch Phys Med Rehab* 1992; **73**: 1063–1068.
- 8 Hamilton BB, Granger CV, Sherwin FS, Zielezny M, Tashman JS. A uniform national data system for medical rehabilitation. In: Fuhrer MJ (ed). *Rehabilitation Outcomes. Analysis and Measurement*. Paul H Brookes Publishing Co., Inc.: Baltimore, MD 1987, pp 137–147.
- 9 Hamilton BB, Laughlin JA, Granger CV, Kayton RM. Interrater agreement of the seven level functional independence measure (FIM). *Arch Phys Med Rehab* 1991; **72**: 790.
- 10 Gresham GE, Labi ML, Dittmar SS, Hicks JT, Joyce SZ, Phillips ML. Quadriplegia index of function. *Arch Phys Med Rehab* 1980; **61**: 493.
- 11 van Tuijl JH, Janssen-Potten YJ, Seele HA. Evaluation of upper extremity motor function tests in tetraplegics. *Spinal Cord* 2002; **40**: 51–64.
- 12 Stroh Wuolle K, van Doren CL, Thrope GB, Keith MW, Peckham PH. Development of a quantitative Hand Grasp and Release Test for patients with tetraplegia using a hand neuroprosthesis. *J Hand Surg* 1994; **19A**: 209–218.
- 13 Mulcahey MJ, Smith BT, Betz RR. Psychometric rigor of the Grasp and Release Test for measuring functional limitation of persons with tetraplegia: a preliminary analysis. *J Spinal Cord Med* 2004; **27**: 41–46.
- 14 Lieshout GV. *User manual Van Lieshout Test*. iRv: Hoensbroek 2003.
- 15 Post MWM, Van Lieshout G, Seelen HAM, Snoek GJ, IJzerman M, Pons C. Measurement properties of the short version of the Van Lieshout test (VLT-SF). *Spinal Cord* [E-pub ahead of Print: 13 June 2006; doi:10.1038/sj.sc.3101937].
- 16 Pfennings LE, van der Ploeg HM, Cohen L, Polman CH. A comparison of responsiveness indices in multiple sclerosis patients. *Qual Life Res* 1999; **8**: 481–489.
- 17 Beaton DE, Bombardier C, Katz JN, Wright JG. A taxonomy for responsiveness. *J Clin Epidemiol* 2001; **54**: 1204–1217.
- 18 Davidson M, Keating JL. A comparison of five low back disability questionnaires: reliability and responsiveness. *Phys Ther* 2002; **82**: 8–24.
- 19 Stratford PW, Binkley FM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther* 1996; **76**: 1109–1123.
- 20 Maynard Jr FM, Bracken MB, Creasey G, Ditunno Jr JF, Donovan WH, Ducker TB et al. International standards for neurological and functional classification of spinal cord injury. American spinal injury association. *Spinal Cord* 1997; **35**: 266–274.
- 21 Post MWM. Pilot-onderzoek Van Lieshout test [in Dutch]. iRv: Hoensbroek 2000.
- 22 Linacre JM, Heinemann AW, Wright BD, Granger CV, Hamilton BB. The structure and stability of the Functional Independence Measure. *Arch Phys Med Rehab* 1994; **75**: 127–132.
- 23 Hall KM, Cohen ME, Wright J, Call M, Werner P. Characteristics of the functional independence measure in traumatic spinal cord injury. *Arch Phys Med Rehab* 1999; **80**: 1471–1476.
- 24 Gresham GE, Labi ML, Dittmar SS, Hicks JT, Joyce SZ, Stehlik MA. The Quadriplegia Index of Function (QIF): sensitivity and reliability demonstrated in a study of thirty quadriplegic patients. *Paraplegia* 1986; **24**: 38–44.
- 25 Marino RJ, Goin JE. Development of a short-form quadriplegia index of function scale. *Spinal Cord* 1999; **37**: 289–296.
- 26 Effing T, Dirx M, Slegers E, Meeteren NV. De responsiviteit van de COPM, ARA en SODA voor verandering in handvaardigheid bij plastisch chirurgische handpatienten. *Ned Tijdschr Ergoth* 1999; **27**: 109–116.
- 27 Wright JG, Young NL. A comparison of different indices of responsiveness. *J Clin Epidemiol* 1997; **50**: 239–246.
- 28 Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990; **28**: 632–642.
- 29 Norman GR, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol* 1997; **50**: 869–879.
- 30 Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989; **27**(Suppl): S178–S189.
- 31 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Academic Press Inc.: New York, NY 1977.
- 32 Heald SL, Riddle DL, Lamb RL. The shoulder pain and disability index: the construct validity and responsiveness of a region-specific disability measure. *Phys Ther* 1997; **77**: 1079–1089.
- 33 Poynton AR, O'Farrell DA, Shannon F, Murray P, McManus F, Walsh MG. Sparing of sensation to pin prick predicts recovery of a motor segment after injury to the spinal cord. *J Bone Joint Surg Br* 1997; **79**: 952–954.
- 34 Marino RJ, Ditunno Jr JF, Donovan WH, Maynard Jr F. Neurologic recovery after traumatic spinal cord injury: data from the Model Spinal Cord Injury Systems. *Arch Phys Med Rehab* 1999; **80**: 1391–1396.
- 35 Ditunno Jr JF, Cohen ME, Hauck WW, Jackson AB, Sipski ML, Stover SL et al. Recovery of upper-extremity strength in complete and incomplete tetraplegia: a multicenter study. *Arch Phys Med Rehab* 2000; **81**: 389–393.
- 36 Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987; **40**: 171–178.