

Clinical Review

Evaluation of upper extremity motor function tests in tetraplegics

JH van Tuijl^{*1}, YJM Janssen-Potten¹ and HAM Seelen^{1,2}

¹Institute for Rehabilitation Research, Hoensbroek, The Netherlands; ²Rehabilitation Foundation Limburg, Hoensbroek, The Netherlands

Objective: To provide an overview of arm–hand function tests useful in tetraplegic subjects. Considerations for selection of an appropriate test are also provided.

Data sources: A Medline literature search was conducted covering the period from 1967 to March 2001. Relevant references cited in the selected papers were also considered, regardless of the year of publication.

Study selection: This review was restricted to strength tests, functional and ADL tests. Only general tests and tests designed specifically to test tetraplegic persons written in English, or in Dutch were included in the review.

Results: Information is provided on four types of strength tests, 10 general and five specific functional tests and eight ADL tests.

Conclusion: Many tests are available to measure upper extremity motor function in tetraplegics. Selection of a test is at first determined by the outcome value in which the investigator is interested. When the type of outcome value has been determined, the most suitable test has to be selected from the range of available tests. When two tests appear to be equally suitable, the availability of information on psychometric properties of the test when used in tetraplegic patients is a decisive factor. When information on the reliability, validity and sensitivity of a test is missing, it should be gathered before using the test.

Sponsorship: The present study was written with financial assistance provided by the Health Research and Development Council of the Netherlands.

Spinal Cord (2002) 40, 51–64. DOI: 10.1038/sj/sc/3101261

Keywords: review, spinal cord injury; tetraplegia; arm–hand function; upper extremity function

Introduction

A spinal cord lesion at the cervical level often results in tetraplegia, with motor, sensory and autonomic function loss. Hanson and colleagues¹ asked tetraplegic persons to choose the function which they would prefer above all others if they could have one function restored. The list of choices included sexual function, bowel and bladder function, walking, and use of arms and hands. Arm–hand function was the most frequent choice.

Therapists agree that restoration of hand function is an important goal in rehabilitation. The main focus in rehabilitation is compensation of function loss, using those parts of the sensorimotor system which are still intact. In some cases muscle/tendon transpositions of intact arm or hand muscles are carried out through reconstructive surgery to substitute lost motor function. Controversy exists among clinicians as to

whether or not these transpositions should be performed in addition to conservative therapy. In both cases patients have to learn new movement strategies in order to perform activities of daily living (ADL). The way in which these new movement strategies develop is still unclear. Recording muscle activity by means of surface electromyography (EMG) may provide insight in the development of these strategies. In order to be able to make inferences about muscle use, a subject should exert several standardised movements during which EMG measurements take place. A functional test could be appropriate to elicit such standardised movements, providing the test fulfils the following criteria: (a) All test items can be performed by tetraplegics with injury levels C4 to C8, and the items are not too long or too strenuous; (b) The test can be administered in a small amount of time; (c) The EMG values resulting from measurements during performance of the tasks are reproducible; (d) The test is sensitive, ie the test detects differences in muscle activity patterns between patients

*Correspondence: JH van Tuijl, Institute for Rehabilitation Research (iRv), PO box 192, 6430 AD, Hoensbroek, The Netherlands

with different levels of spinal cord injury (SCI) and within patients over time; (e) The test resembles functional tasks; (f) The test elicits several movements in the shoulder, and elbow joint, and of the lower arm, instead of only one movement, as in tests in which the same movement has to be performed several times; (g) The test provides movements with a clear defined start and endpoint.

When searching for literature on health status assessment, Deyo and Patrick² found that finding relevant literature is often difficult, given the variety of labels attached to health status instruments, the paucity of search terms in computerised databases, and the dispersion of relevant articles throughout the literature of many disciplines. The same problems were encountered during the search for relevant publications focusing on arm-hand function tests in tetraplegic persons.

Although many descriptions of tests designed to evaluate upper extremity function are available, only a few seem suitable for use in tetraplegic persons. Furthermore, it is difficult to decide which of the available tests is superior, because they often have more or less the same content and purpose, and because information on reliability and validity has not always been published.

In view of the above mentioned, a critical review which summarises the characteristics of the available upper extremity tests is required. A review of functional hand evaluations was published in 1987 by McPhee.³ Wade⁴ reviewed tests designed to measure arm impairment and disability in stroke patients. Reviews that focused on upper extremity function tests in tetraplegics included only references to the available tests, but did not provide a critical description of the tests.^{5,6}

The purpose of the present paper is to give an in depth overview of upper extremity tests that can be used during evaluation of arm-hand function in tetraplegics. The present review focuses on strength tests, functional tests and ADL tests. Sensibility and exercise tolerance are aspects of arm-hand function, often included in the previously mentioned tests, which are not discussed separately.

Methods

A Medline literature search was conducted covering the period from 1967 to March 2001 using combinations of the key words 'tetraplegia', 'tetraplegic(s)', 'quadriplegia', 'paraplegia', 'spinal cord injury/injuries', 'SCI', 'paresis', 'assessment', 'test', 'index', 'evaluation', 'function', 'strength', 'dexterity', 'skill(s)', 'hand', 'upper extremity', 'motor', 'activities-of-daily-living', and 'ADL'. In addition, the references cited in the selected papers were considered, regardless of the year of publication. Only publications written in English or Dutch were included in this review.

In the present paper a description of functional and ADL tests is given according to the following items:

purpose of the test, target population, composition of the test and scoring method. When available, information on reliability, validity and sensitivity is summarised. When normative data are available a reference is provided.

The upper extremity motor function tests are classified in the following categories: (1) Strength tests; (2) Functional tests; (3) ADL tests.

In this paper two categories of functional tests are discussed, namely (a) general functional tests, which have been designed for a broad category of patients and (b) specific hand function tests, which were designed to evaluate tetraplegic persons. Most functional tests primarily measure the performance of specific tasks under standardised conditions. These tasks are more abstract than the tasks, which are performed in the so-called ADL tests. ADL tests score the ability to perform certain ADL activities in a standardised situation.

Most tests included in this review fit into the above mentioned classification system, except for the Sollerman test,⁷ which measures ADL but is known as a hand function test. This test was classified as a functional test. The Action Research Arm (ARA) test, which is solely used in stroke patients, does not fulfil the inclusion criteria as mentioned above. However, because this test is based on the Upper Extremity Function Test (UEFT), a general functional test, the ARA was, nevertheless, included in the present review.

Results

Strength tests

Methods used to measure strength include manual muscle testing (MMT), hand-held dynamometry, pinch and grip strength measurement, and isokinetic dynamometry.

Manual muscle testing (MMT) When using MMT an examiner counteracts the force of a subject manually. The extent to which the subject is capable to counteract the examiner's force is recorded. The 6-level scheme proposed by the Medical Research Council⁸ is often used as a scoring scale for MMT (see Table 1).

When determining the ASIA score,⁹ key muscles are tested using MMT. The category 'not testable' (NT) was added to the MRC scale. A disadvantage of using MMT in muscles with impaired innervation is that a maximum MRC score of 2 can be reached, which

Table 1 MRC grades⁸

0	= No contraction
1	= Flicker or trace of contraction
2	= Active movement, with gravity eliminated
3	= Active movement against gravity
4	= Active movement against gravity and resistance
5	= Normal power

reduces the number of scoring categories. Reliability of MMT in SCI persons has not been determined yet. According to Noreau *et al*¹⁰ MMT is not sufficiently sensitive to assess muscle strength, at least for grade 4 and higher and to detect small or moderate increases of strength in SCI persons over the course of rehabilitation. On the other hand, Waters *et al*¹¹ state that in tetraplegics (injury level C3 to C7) MMT is sensitive enough to detect changes in strength of key muscles in time. Because they only tested the key muscles, as indicated by the ASIA classification, most MRC scores were in the 0–3 range, whereas the limitations of MMT mentioned by Noreau *et al* (1998) concern in particular MMT scores of 4 and 5.

Hand-held dynamometry Several hand-held dynamometers have been used to test muscle strength in tetraplegics, for example a Penny and Giles dynamometer.^{10,12} Only muscles with a minimum MMT score of 3.5 can be tested with hand-held dynamometry, because it is not possible to test a muscle with lower MMT scores, unless the movement is made in a position in which the influence of gravity on the movement is minimised.¹³ May *et al*¹⁴ emphasised the difference between a break test (the investigator overcomes the strength of the tested person) and a make test (the investigator holds the dynamometer stationary while the subject exerts a maximal strength against it). They used a break test to measure isometric strength of shoulder rotation in SCI persons. This resulted in values with good intrarater reliability ($r=0.89$ to 0.96). Interrater reliability of dynamometry of the extensor carpi radialis muscle in tetraplegic persons is good (ICC=0.83).¹⁵ Drolet *et al*¹⁶ used hand-held dynamometry to evaluate the strength of six muscle groups in both paraplegic and tetraplegic subjects. Other investigators^{10,12,13,15} tested SCI persons both with MMT and hand-held dynamometry. They all emphasised the importance of hand-held dynamometry as a useful supplement to the MMT method. Use of a hand-held dynamometer may identify effects of therapeutic interventions, missed by MMT, especially for MMT grades 4 to 5. Bohannon¹⁷ provided reference values for upper extremity muscle strength using hand-held dynamometry in healthy subjects, as well as statistics on test–retest reliability.

Grip and pinch strength measurement Several devices are available to measure grip strength. Mathiowetz *et al*¹⁸ provide recommendations about a standard test protocol, in which they recommend the Jamar dynamometer to measure grip strength. Information on reliability and validity of grip and pinch strength evaluations in healthy women is available.¹⁸ Normative values for grip and pinch strength in adults¹⁹ and for grip strength in elderly²⁰ are available. Pinch dynamometry appears to be useful to measure improvement in grip strength after hand-surgery in tetraplegics.²¹

Several other investigators did not use standard available equipment to measure grip and pinch

strength, but developed devices themselves, which can be used to measure strength in SCI persons.^{22–24} Some of the functional tests mentioned below also include grip strength measurements.^{25,26}

Isokinetic dynamometry Most investigators consider isokinetic dynamometry as the ‘gold standard’ to assess muscle strength.¹⁰ Its clinical use, however, is limited because of the expensive equipment, the dimension of the apparatus, the time required for the subject’s positioning and the assessment procedures. Furthermore a MRC grade of at least 3 is necessary to perform the desired movement, whereas paretic muscles (MRC \leq 2) cannot overcome gravity and therefore cannot move the dynamometer over the entire range of motion required to test the muscle.

May *et al*¹⁴ measured shoulder strength of SCI persons with both hand-held and isokinetic dynamometry. The correlation between both measures was high (Pearson product moment correlations of 0.86 and 0.88). Because of the high correlation one can conclude that the hand-held dynamometer is a good alternative to measure (shoulder) strength in SCI persons.

Van der Ploeg and Oosterhuis²⁷ described advantages and disadvantages of MMT, hand-held dynamometry and isokinetic dynamometry. For more information on normative values and the influence of factors like age, gender, upper extremity position, and handedness, the reader is advised to consult the abundant available publications on this matter.^{10,12,13,17,20,27,28}

Functional tests

Contrarily to publications on hand function in rheumatoid arthritis (RA) and stroke patients, literature on functional tests used in tetraplegic persons is sparse. In the next paragraph several general and more specific functional tests will be discussed. The general tests, which will be discussed below, were either developed to measure function in a broad patient group or were initially designed to measure function in a specific population, but were later used in other populations including tetraplegics. The disease-specific functional tests mentioned in this paper were all designed to measure hand function in tetraplegic persons. The following functional tests are discussed below:

General functional tests:

- – the Minnesota Rate of Manipulation (MRM) test
- – the Upper Extremity Function Test (UEFT)
- – the Purdue Pegboard test
- – the Jebsen test of hand function
- – the Nine-Hole Peg test
- – the Smith hand function evaluation
- – the Box and Block Test (BBT)
- – the Physical Capacities Evaluation of Hand Skill (PCE)
- – the Action Research Arm (ARA) test
- – the Sollerman hand function test

Tests specifically designed for tetraplegic persons:

- – the Standardised Object Test (SOT)
- – the Vandenberg hand function test
- – the Grasp and Release Test (GRT)
- – the Capabilities of Upper Extremity (CUE) Instrument
- – Thorson's functional test

General functional tests

Name: the Minnesota Rate of Manipulation (MRM) test.

Investigator(s): Fleishman (1964).²⁹

Purpose: measurement of manual dexterity.

Target population: the test was initially designed to test healthy subjects, but in the course of time the test has also been used in rehabilitation settings.

Test composition: five subtests, including placing, turning, displacing, one-hand turning and placing, and two-hand turning and placing.

Scoring method: the time necessary to perform each subtest.

Psychometric properties: the interrater reliability of the MRM test was good ($r=0.75$).³⁰ The scores on four subtests (the turning subtest was excluded) of the MRM test were compared to the American Medical Association's (AMA) Rating Scale, which lead to the conclusion that the MRM test provides a more limited but better-defined assessment of hand impairment than the AMA Rating Scale.³⁰

Normative data: means and standard deviations on MRM test scores for subjects with impaired hand function are available.³⁰

Additional information: The original manuscript in which the test is fully described, was not considered in this paper.

Name: the Upper Extremity Function Test (UEFT).

Investigator(s): Carroll (1965).³¹

Purpose: provide a semi-quantitative test of upper extremity function.

Target population: patients with upper extremity impairment.

Test composition: the test consists of several tasks, including moving objects to a shelf, placing them over a peg, writing one's name, placing the hand to mouth, head, and neck, and pouring water from a pitcher or glass. The objects are of different shapes and weights designed to test grasp, grip, pinch, placing, arm extension and elevation, pronation and supination, and, to a lesser extent, strength. A wooden box is necessary to serve as a shelf and the objects to be handled have to be available. Administration of the test takes approximately 1 h.³²

Scoring method: the items are graded on a four-point scale (0–3). Maximum score for the right hand is 99, and for the left (non-writing) hand 96.

Psychometric properties: information on test–retest and interrater reliability as presented by Carroll³¹

suggested good reliability, but no statistical analysis of the results was performed. The test was validated by correlating the UEFT to a standard hand activities of daily living test. Only a scatter plot without statistics was provided on these results.³¹ Lyle³² reproduced Carroll's test and found high correlations (0.98 for both the impaired and non-impaired side) when assessing test–retest reliability. Interrater reliability was also high ($r=0.99$).

Normative data: not available.

Additional information: a summary description of the use of the UEFT in tetraplegics is available.³³

Name: the Purdue Pegboard test.

Investigator(s): Tiffin (1948).³⁴

Purpose: measurement of unilateral and bilateral fine manual dexterity.

Target population: originally the test was developed to select personnel in industries. In the course of time the test has also been used in rehabilitation settings.

Test composition: five subtests in which prehension of very small pins, washers, and collars with the right hand (RH), left hand (LH), both hands (BH), right + left + both (R + L + B) hands and an assembly subtest are scored. The assembly subtest requires that both hands work simultaneously while performing different tasks for 60 s.

Scoring method: for the RH, LH and BH subtests the number of pegs placed during 30 s is recorded. The R + L + B subscore can be calculated from the scores of the first three tests. The score for the assembly subtest is the total number of pins, washers, and collars placed in 60 s.

Psychometric properties: ICC values for test–retest reliability of each subtest of the Purdue Pegboard test ranged from 0.37 to 0.70 during one-trial administration and from 0.81 to 0.89 for the average of three trials.³⁵ Scores seemed to be more reliable when three trials for each subtest were averaged. In accordance with this finding it was suggested that therapists who administer the one-trial test should exercise caution when interpreting improved scores.³⁶ The examiner manual published in 1968³⁷ is based on a revised version of the Purdue Pegboard test. Therefore, it is recommended that only reliability values obtained while using the revised version of the Purdue Pegboard test should be used.

Normative data: normative values for male and female healthy subjects,³⁸ 14- to 19-year-old healthy children,³⁹ and healthy elderly³⁶ are available.

Name: the Jebsen test of hand function.

Investigator(s): Jebsen et al (1969).⁴⁰

Purpose: assessment of hand disability and improvement in hand function gained by therapeutic procedures.

Target population: patients with hand disabilities, including hemiparetic patients, rheumatoid arthritis patients, and patients with C6–7 traumatic tetraplegia.⁴⁰ Other authors used the test to assess function in

hemiplegic patients,⁴¹ tetraplegic adults,⁴² a tetraplegic child,⁴³ and children with cerebral palsy.⁴⁴

Test composition: seven unilateral subtests, including writing, turning over cards, picking up small common objects, simulated feeding, stacking checkers, picking up large light objects, and picking up large heavy objects. In healthy subjects, both hands can be tested in approximately 15 min.

Scoring method: time necessary to complete each subtest.

Psychometric properties: in the dominant hand the test-retest reliability in patients with stable hand disorders was high ($r=0.89$ to 0.99) except for the writing subtest ($r=0.67$). With the non-dominant hand the simulated feeding subtest was the least reproducible ($r=0.60$). The test seemed to be sensitive enough to detect changes in hand function in patients with different impairments.⁴⁰

Normative data: mean times and standard deviations of healthy subjects are available.^{40,45}

Name: the Nine-Hole Peg test.

Investigator(s): Kellor (1971).⁴⁶

Purpose: measurement of dexterity.

Target population: healthy subjects and persons with impaired dexterity.

Test composition: unilateral test in which nine pegs have to be placed in a board and then removed.

Scoring method: time necessary to perform the task.

Psychometric properties: the interrater reliability of the Nine-Hole Peg test in healthy subjects was good (right hand $r=0.97$, left hand $r=0.99$). The test-retest reliability was moderate (right hand $r=0.69$, left hand $r=0.43$). Concurrent validity of the Nine-Hole Peg in healthy subjects test was determined by comparison with the Purdue Pegboard test. This resulted in significant correlations (right hand $r=-0.61$, left hand $r=-0.53$).⁴⁷

Normative data: normative values for healthy males and females are available.^{46,47}

Name: the Smith hand function evaluation.

Investigator(s): Smith (1973).²⁶

Purpose: provide a standardised evaluation of hand co-ordination and functional hand skills associated with activities of daily living.

Target population: patients having hand dysfunction problems, that is, inco-ordination, poor muscle strength, decreased sensation, and limited range of motion.

Test composition: four subtests, including unilateral grasp-release tasks, activities of daily living, a writing ability task and a grip strength measurement. A total of 13 items are tested. Because many patients with severe hand dysfunction problems are slowed down considerably by their disabilities, testing time may take as long as 45 to 60 min or more, depending upon type of impairment and degree of involvement of the upper extremity.

Scoring method: in most sections of the evaluation time is measured, in the dynamometer subtest force is measured.

Psychometric properties: not available.

Normative data: normative values for healthy subjects are available.²⁶

Name: the Box and Block Test (BBT).

Investigator(s): Cromwell and colleagues (1976).⁴⁸

Purpose: providing a basic measure of gross manual dexterity which can be easily and quickly given.

Target population: healthy and handicapped individuals.

Test composition: unilateral test in which wooden blocks have to be transported from one compartment to another.

Scoring method: the number of blocks transported from one compartment into another during 1 min is recorded. Each trial is preceded by a 15 s trial period.⁴⁹

Psychometric properties: test-retest reliability was high (ρ coefficients of 0.937 and 0.976 for the left and right hand respectively),⁴⁸ and interrater reliability in healthy subjects was also high (right hand $r=1.0$, left hand $r=0.99$).⁴⁹ Validity of the BBT in elderly was determined by comparison to the ARA test and the Functional Autonomy Measurement System (SMAF).⁵⁰ The BBT appeared to be more related to an independence measurement (the SMAF) than to the ARA test. Comparison of the BBT to the Minnesota Rate of Manipulation (MRM) placing test resulted in high correlations ($r=0.91$) when testing handicapped adults.⁴⁸

Normative data: normative values of adults,⁴⁹ adults with neuromuscular conditions,⁴⁸ and healthy elderly⁵⁰ are available.

Name: the Physical Capacities Evaluation of Hand Skill (PCE).

Investigator(s): Bell et al (1976).²⁵

Purpose: provide an objective measurement of performance.

Target population: paraplegic, tetraplegic and hemiplegic persons. Tetraplegics can be tested both with and without an orthosis.

Test composition: five unilateral hand skill tests, seven bilateral hand skill tests, and a dynamometer reading.

Scoring method: the tests are timed for either a specified period or for completion of the task. Test scores were standardised to normative values derived from the results of 50 healthy controls. Graphic comparison of the results is possible.

Psychometric properties: not available.

Normative data: not available.

Name: the Action Research Arm (ARA) test.

Investigator(s): Lyle (1981).³²

Purpose: provide a rapid yet reliable and standardised performance test appropriate for use in assessing recovery of upper limb function following cortical damage.

Target population: hemiplegic persons.

Test composition: four subtests, including grasp, grip, pinch and gross movement. A total of 19 items are scored. In the Dutch version of the ARA test⁵¹ the subtest pinch was divided in a subtest gross pinch and fine pinch.

Scoring method: the items are graded on a four-point scale (0–3). In the Dutch version time limits for each item were added.⁵²

Psychometric properties: the ARA has good intrarater, interrater and test–retest reliability in stroke patients (all Spearman correlations >0.98).^{52,53} The concurrent validity of the ARA has been confirmed by comparison with the Brunnstrom-Fugl-Meyer Assessment ($r > 0.91$),^{54,55} with the Sollerman test ($r = 0.94$),⁵² and with the motor assessment scale (MAS) (ICC = 0.98).⁵⁶

Normative data: not available.

Additional information: the test is based on Carroll's Upper Extremity Function Test (UEFT).³¹

Name: the Sollerman hand function test.

Investigator(s): Sollerman and Ejeskär (1995).⁷

Purpose: giving a good measure of overall function of the hand.

Target population: tetraplegics, rheumatoid arthritis patient, finger amputees, nerve injured persons, persons with impaired range of motion of the arm.

Test composition: 17 unilateral and three bilateral (total 20) activities of daily living. As the upper time limit for each subtest is 1 min, the test can usually be completed within 20 min.

Scoring method: each subtest is scored on a five-point scale (0–4), with a maximum score of 80 points for the dominant hand and 77–79 points for the non-dominant hand.

Psychometric properties: interrater reliability was good ($r = 0.98$). The test score correlated well with the accepted international functional classification of the patient's arm ($r = 0.76$, $P < 0.001$). The mean test score in the arms of patients lacking sensation was significantly lower than in those with tactile gnosis (O:1–3 compared with OCu:1–3, $P < 0.001$).

Normative data: mean Sollerman test scores for each functional group according to the International classification for surgery of the hand in tetraplegia are provided.⁷

Additional information: Curtin⁵ comments that this test is only suitable for those persons who have some wrist movement.

Functional tests for tetraplegic persons

Name: the Standardized Object Test (SOT).

Investigator(s): Thrope et al (1989).⁵⁷

Purpose: evaluation of the minimal criteria of functional hand grasp necessary to use a FNS neuroprosthetic hand system.

Target population: tetraplegic persons.

Test composition: the test consists of six objects each having various weights, sizes, and textures, including a

block, disk, videotape, pegs, cylinder, and fork. The subject is asked to acquire, transport, and release each object as many times as possible in a 30-s period.

Scoring method: number of objects transported.

Psychometric properties: the test was sensitive enough to detect an increase in hand function in tetraplegics when using a hand system.

Normative data: not available.

Additional information: the same group also developed the Common Object Test⁵⁸ and the Grasp and Release Test.⁵⁹

Name: the VandenBerge hand and arm function test.

Investigator(s): VandenBerge et al (1991).²¹

Purpose: evaluation of the effect of reconstruction surgery.

Target population: tetraplegic persons.

Test composition: nine unilateral items, including transfer of bowls of different weights, grasp and transfer of different objects, and writing a sentence. The duration of the test is dependent on the speed with which a subject performs the subtests.

Scoring method: time necessary to perform each subtest.

Psychometric properties: not available.

Normative data: mean times necessary to perform each subtest for 13 tetraplegics were reported without distinguishing between subjects with different injury levels.

Name: the quantitative hand Grasp and Release Test (GRT).

Investigator(s): Stroh Wuolle et al (1994).⁵⁹

Purpose: assessing the use of a hand neuroprosthesis in C5 and C6 level tetraplegic persons.

Target population: tetraplegic persons.

Test composition: subjects grasp, move, and release one of six different objects as many times as possible in five 30-s trials for each object, with and without a neuroprosthesis. Three objects have to be manipulated with lateral prehension (peg, paperweight, and fork) and three with palmar prehension (block, can, and videotape).

Scoring method: the number of completions and failures within each 30-s trial. Each subtest includes a pretest and five attempts to transport as many objects as possible.

Psychometric properties: the reproducibility and validity of the test were not determined and Stroh and colleagues tested only five subjects of whom only the results of four subjects, two C5 and two C6, were reported.

Normative data: not available.

Additional information: This test is very similar to the SOT reported on by Thrope, Stroh and Baco in 1989.⁵⁷ Smith et al⁶⁰ used the GRT to evaluate the effect of FNS in tetraplegic adolescents. They concluded that the assessment was suitable for use with adolescents as there were no apparent age-related differences in performance.

Name: the Capabilities of Upper Extremity (CUE) instrument.

Investigator(s): Marino et al (1998).⁶¹

Purpose: measurement of upper extremity functional limitations in individuals with tetraplegia.

Target population: tetraplegic persons.

Test composition: 32 items have to be scored by an interviewer during a telephone interview.

Scoring method: Responses are given on a seven-point scale representing self-perceived difficulty in performing the action, varying from (1) unable to perform and (7) can perform without difficulty.

Psychometric properties: the Cronbach's α of the test was 0.96 and the test-retest reliability was good (ICC=0.94). In general, individuals with more caudal motor levels received higher scores on the partial CUE. *Post hoc* pairwise comparisons indicated that the mean CUE scores were significantly different for motor levels more than one level apart except for C7 versus T1 on the right side. The CUE was less successful at differentiating adjacent motor levels. The CUE displayed a high correlation with the upper extremity motor score (UEMS) and the motor Functional Independence Measure (FIM) score.

Normative data: Mean CUE values are provided for tetraplegic persons with different levels of injury and by best motor level.

Name: Thorson's functional test.

Investigator(s): Thorson et al (1999).⁶²

Purpose: evaluation of hand function when using a stimulation device, the MeCFES.

Target population: tetraplegic persons.

Test composition: eight unilateral tasks which are divided into four groups, including moving flat objects, namely CD covers of different weights and a thin book, moving cylindrical objects, drinking, and eating with a spoon. The total experiment, including preparation, takes less than 1.5 h.

Scoring method: the performance of the grip is rated on a three-point scale (0–2).

Psychometric properties: not available.

Normative data: not available.

ADL tests

In evaluation studies of upper extremity function in tetraplegic persons, not only functional tests are used, but also tests in which the subject is asked to perform several activities of daily living (ADL tests). For a tetraplegic person an improvement in ADL is often more meaningful than an improvement in the ability to move objects, which is what is measured by many functional tests. Several ADL tests are currently in use, however, most of them are not appropriate to use in tetraplegics, either because they test functions, which are in general not disturbed in tetraplegics, for example communication, and/or because they include items, which cannot be performed by tetraplegics, like for example walking. In the present review only tests which

have been used in spinal cord injured persons were included. The following ADL tests are discussed below:

- two adapted versions of the Barthel Index (BI)
- the Functional Independence Measure (FIM)
- the Ranchos Los Amigos Hospital (RLAH) Functional Activities test
- the Quadriplegia Index of Function (QIF)
- the Common Object Test (COT)
- the Spinal Cord Independence Measure (SCIM)
- the 'Valutazione Funzionale Mielolesi' (VFM)

Name: Dutch interview version of the Barthel Index (iv-BI).

Investigator(s): Post et al (1995).⁶³

Purpose: assessing ability to cope in activities of daily living by SCI persons.

Target population: spinal cord injured persons.

Test composition: 10 items, including personal care, toileting, bladder and bowel management, eating, transfers, ambulation, dressing, stair climbing and bathing.

Scoring method: items are scored on two to four point scales (0–1 to 0–3), a maximum score of 20 can be obtained, in which a higher score implies greater independence.

Psychometric properties: the interview version of the BI appeared to be a reliable test to measure ADL independence in SCI persons (Cronbach α : 0.87). In persons with complete SCI a strong correlation (Spearman-correlation: 0.69; $P < 0.0001$) between level of injury and adapted BI-scores existed. The mean iv-BI-score of complete tetraplegics was significantly lower than the scores in incomplete tetraplegics and paraplegics. The iv-BI was also sensitive enough to differentiate between a C3,4,5 and a C7,8 level group and between a C6 and a C7,8 group. The iv-BI was unable to differentiate between the C3,4,5 and C6 group.

Normative data: iv-BI scores are provided for complete and incomplete tetraplegics and by level of injury.

Additional information: Post et al⁶³ translated Collins version of the BI in Dutch and made it suitable as a patient questionnaire. Collins version of the BI⁶⁴ is a slightly modified version of the original BI⁶⁵ in which mainly the scoring system was adapted.

Name: modified Barthel Index (MBI).

Investigator(s): Granger et al (1979),⁶⁶ and Yarcony et al (1987).^{67,68}

Purpose: measurement of severity of disability and monitoring of rehabilitation progress in severely disabled persons,⁶⁶ or assessment of functional abilities.⁶⁸

Target population: traumatic SCI persons.

Test composition: the MBI consists of 15 tasks,⁶⁸ including drinking from a cup, feeding from a dish, upper body dressing, lower body dressing, donning a brace or prosthesis, bathing, grooming, bowel continence, bladder continence, chair transfers, toilet transfers, tub/shower transfers, walking, stair-climbing, and wheelchair propulsion (only if not walking).

In Yarcony's investigation published in 1987 the item 'donning brace or prosthesis' was not included.

Scoring method: items are rated as independent, assisted, or dependent. Items that are considered more important for independence, such as eating without assistance, are weighed more heavily than less important items, like grooming.

Psychometric properties: the MBI was able to identify statistically significant improvement from discharge to 3-year follow-up in both complete and incomplete tetraplegics.⁶⁸

Normative data: self-care and mobility subscores of the MBI at admission and discharge for patients with complete and incomplete tetraplegia are provided,⁶⁷ as are mean MBI scores during 3-year follow-up.⁶⁸

Name: the Functional Independence Measure (FIM).

Investigator(s): Hamilton *et al* (1987 and 1991).^{69,70}

Purpose: rating severity of patient disability and the outcomes of medical rehabilitation.

Target population: patients who undergo medical rehabilitation.

Test composition: 18 items, concerning self-care (eating, grooming, bathing, dressing upper body, dressing lower body, toileting), sphincter control (bladder and bowel management), mobility (transfers to bed, chair or wheelchair, to toilet, and to tub or shower), locomotion (walking or wheelchair propulsion, stair climbing), communication (comprehension and expression), and social cognition (social interaction, problem solving, memory).

Scoring method: the items are scored on a seven-point scale, varying from (1) total assistance to (7) complete independence.

Psychometric properties: the FIM appeared to have good clinical interrater agreement in patients undergoing inpatient medical rehabilitation (ICC=0.97).⁷⁰ FIM scores were significantly lower in complete C4 tetraplegics than in C6 tetraplegics,⁷¹ which indicated that the FIM is sensitive enough to differentiate between different levels of injury. In incomplete tetraplegic persons FIM scores appeared to change significantly between admission and discharge. In complete tetraplegics no significant change was found.⁷² Hall *et al*⁷³ found improvements in FIM score in each injury level group (C1–C3 to C8). These groups were not divided according to completeness of the injury. The former mentioned results indicate that the FIM is useful in detecting changes in function in time. FIM motor gains were greatest between admission and discharge for all neurologic levels. Because the FIM was never intended for use with outpatients or at subacute levels of care, this is no disadvantage of the FIM.⁷³ FIM scores appeared to reach a plateau in most SCI persons.^{74,75} Hall *et al*⁷³ defined scores 6 and 7 as 'ceiling' score, and a score of 1 as 'floor' score. High tetraplegics (C1–C4) appeared to have 86% floor scores at admission on the motor FIM. This percentage was 21% at discharge. In low tetraplegics (C5–C8) 61% floor scores were obtained

at admission, decreasing to 3% at discharge. For the cognitive FIM scores 59 and 67% ceiling scores were obtained at admission and these percentages increased to 80 and 86% at discharge in high and low tetraplegics respectively. This might indicate that the motor items of the FIM are too difficult for tetraplegics and the cognitive items measure aspects which are not affected by the spinal cord lesion.

Normative data: mean FIM scores by injury level and age,⁷¹ and by injury level and Frankel grade over time⁷³ are available. Caution has to be paid when comparing the FIM score of an individual patient to these norms, because several factors may influence the FIM score, namely age,⁷¹ length of stay, and level of education. In low (C1–C4) tetraplegics the rate of change of FIM scores between admission and discharge was not related to length of stay and age. But, older patients did tend to reach a lower level of plateau. In high (C5–C8) tetraplegics a small effect of level of education on rate of change was noted, in which a higher level of education was associated with a more rapid rate of recovery.⁷⁵ Warschauski *et al*⁷⁵ noticed an effect of length of stay on FIM motor scores in paraplegics, but not in tetraplegics.

Additional information: The first version of the FIM used a four-point rating scale (0–4) to score each item.⁶⁹ A revised version of the FIM has been developed, which uses the above-mentioned seven-point scale.

Name: the Ranchos Los Amigos Hospital (RLAH) Functional Activities test.

Investigator(s): Rogers and Figone (1980).⁷⁶

Purpose: assessment of self-care skills in tetraplegics.

Target population: tetraplegic persons.

Test composition: eight categories are included, namely feeding, grooming, toileting and bathing, upper extremity dressing, lower extremity dressing, written communication, desk skills and transfers. Three to seven items are tested within each category.

Scoring method: the items are rated on a three-point scale, namely independent, assisted or unable. The test also assesses the use of upper extremity orthotic and assistive devices.

Psychometric properties: not available.

Normative data: not available.

Name: the Quadriplegia Index of Function (QIF).

Investigator(s): Gresham *et al* (1980).⁷⁷

Purpose: provide a more specific and sensitive instrument to document the functional improvements achieved during the rehabilitation of tetraplegic patients.

Target population: tetraplegic persons.

Test composition: the index is composed of 10 variables: transfers, grooming, bathing, feeding, dressing, wheelchair mobility, bed activities, bladder program, bowel program, and understanding personal care. Administration of the test takes 30 min or less.

Scoring method: the items are graded on a five-point scale (0–4) in order of increasing independence.

Psychometric properties: the interrater reliability of the QIF was good (Pearson's $r=0.68$ to 0.98).⁷⁸ The QIF appeared to improve significantly in both complete and incomplete tetraplegics between admission to and discharge from medical rehabilitation.^{79,80} Comparison of the total QIF to the total FIM resulted in a high correlation ($r=0.97$).⁸⁰ Comparison of subgroups of the QIF and FIM also resulted in high correlations between the subtests, except for the feeding subtest.⁸¹ The QIF seemed to assess functional ability in the category of feeding more accurately than the FIM.

Normative data: average scores on the QIF at admission and discharge are provided for persons with complete and incomplete tetraplegia.⁸⁰

Additional information: in 1999 Marino and Goin⁸² developed a short-form version of the QIF (sf-QIF). The sf-QIF consists of six items, and is also graded on a five-point scale. The following items were selected: wash/dry hair, turn supine to side in bed, put on lower body clothing, open carton/jar, transfer from bed to chair, and lock wheelchair. Contrary to the original QIF the individual items in the sf-QIF were not weighted when determining the total score. The sf-QIF score and the 37-item QIF score correlated highly (Spearman correlation = 0.978).

Name: the Common Object Test (COT).

Investigator(s): Stroh et al (1989).⁵⁸

Purpose: evaluation of the use of functional nerve stimulation (FNS).

Target population: tetraplegic persons.

Test composition: the COT uses a task analysis approach to evaluate a person's ability to perform specific phases of an activity. Each ADL is broken down into phases, including acquire and release phases and several performance phases unique to each activity. For example, the performance phases of eating are stab, lift-lower, and bite.

Scoring method: the subject is scored on (1) independence of performance; (2) quality of performance; (3) preference; (4) frequency of an activity; (5) frequency of method; (6) frequency of method at the observed level of independence for both systems; and (7) importance of the activity to the subject. The scoring of independence of performance, ie physical assist, adaptive equipment, self-assist, or independent, is assigned for each phase of the activities.

Psychometric properties: not available.

Normative data: not available.

Additional information: Mulcahey et al⁸³ also used this test to evaluate FES in adolescents with C5 or C6 level tetraplegia.

Name: the Spinal Cord Independence Measure (SCIM).

Investigator(s): Catz et al (1997).⁸⁴

Purpose: disability scale developed specifically for SCI persons in order to make the functional assessments of persons with paraplegia or tetraplegia more sensitive to changes.

Target population: persons with spinal cord injury.

Test composition: the SCIM covers three areas of function: self-care (score range 0–20), respiration and sphincter management (0–40), and mobility (0–40). The time needed for the evaluation is 30 to 45 min.

Scoring method: 16 items are scored on an ordinal scale varying from three to nine classes. The final score ranges between 0 and 100.

Psychometric properties: the interrater reliability of the total SCIM scores was good ($r=0.98$). Sensitivity of the SCIM appeared to be higher than the sensitivity of the FIM. In tetraplegic subjects the FIM missed 22% of the functional changes detected by the SCIM.⁸⁵

Normative data: not available.

Additional information: in 2001 the developers of the SCIM presented a revised SCIM. The interrater reliability of total revised SCIM scores was also high ($r=0.99$, $P<0.0001$). The correlation of total scores by the Catz-Itzkovich SCIM and the total FIM scores is high (Spearman correlation = 0.84 , $P<0.001$). No distinction was made between paraplegics and tetraplegics. Therefore, no recommendations can be made on the usability of the revised SCIM in tetraplegics. When the SCIM is chosen as evaluation tool, the authors recommend the use of the revised SCIM.⁸⁶

Name: the 'Valutazione Funzionale Mielolesi' (VFM).

Investigator(s): Taricco et al (2000).⁸⁷

Purpose: description and assessment of functional status in persons with SCI.

Target population: spinal cord injured persons.

Test composition: the test consists of 65 items, divided in eight general domains of ADL, namely bed mobility, feeding, transfers, wheelchair use, grooming and bathing, dressing, social vocational skills, and standing and walking. The last item was not mentioned in the appendix in which the item-level distribution on VFM scores at baseline were reported. The test takes 30 to 50 min.

Scoring method: the items are reported on a five-point scale (1–5).

Psychometric properties: according to the authors the test appeared to be a reliable and valid instrument for assessing functional status of persons with SCI. Unfortunately, information about the interrater reliability and applicability of the VFM has been reported in Italian.

Normative data: not available in English or Dutch.

Additional information: the authors emphasise that the VFM should not be used for normative assessments of a patient's functional status. Rather, the VFM's main potential use is to monitor individual patients over time, document changes, and assess treatment effectiveness. Moreover, results of patient performances at each domain can lead to the identification of areas where patients need specific training. VFM should be considered essentially as a working tool for daily patient care in rehabilitation

and not as a research-oriented instrument where priority should be given to shorter and simpler scales.

Discussion

The purpose of the present paper was to give an extensive overview of tests, which may be used to evaluate the effect of an intervention on upper extremity function in tetraplegic persons. Before the 1980s few upper extremity tests useful for measuring function in tetraplegic persons were available, with the exception of the Physical Capacities Evaluation of Hand Skill (PCE).²⁵ Recently, more attention has been paid to arm-hand function tests, mainly due to interesting developments in reconstructive surgery and functional electrical stimulation (FES).

When selecting an evaluation instrument several considerations have to be made. First of all, one has to determine what the main purpose of the intervention under investigation is. Is the intervention directed at for example increasing grip strength, or at improving ADL independence, or are several goals pursued. Dependent on the purpose of the intervention under study a suitable outcome measure or several outcome measures should be selected.

When strength is chosen as an outcome value, hand-held dynamometry seems to be a good method, with higher sensitivity than MMT in the range of MRC grades 4 and 5, and without the disadvantages associated with isokinetic dynamometry. Interventions directed at improving grip strength may be evaluated by grip and pinch dynamometry.

Performance during functional activities as an outcome parameter requires selection of an appropriate functional test. Several considerations influence the choice of the test. Firstly, one has to choose between a standard available functional test and a test developed by the investigators own group. Using widely available functional tests facilitates comparison between studies in which different interventions are evaluated with the same evaluation tests. In view of the fact that developments in reconstructive surgery and functional electrical stimulation evolve in centres all over the world, this is a considerable advantage. The advantage of using a self-developed test is that the test is focused on those aspects of upper extremity function which are expected to change due to the intervention. A disadvantage of most self-developed tests, however, is that the reliability, validity and sensitivity often have not been determined.

Secondly, one has to choose between a disease-specific and a general test. Advantages of a disease-specific instrument are a better focus on functional areas of particular concern, and perhaps greater responsiveness to disease-specific interventions. As mentioned above, an advantage of a general test is that it permits comparisons across interventions and diagnostic conditions.² Stroh Wuolle and colleagues⁵⁹ used a general test, the Jebsen test, to evaluate hand function in tetraplegics with and without a neuro-

prosthesis. Some of the problems they encountered during administration of the Jebsen test are illustrative of the disadvantages of using a general functional test. The following problems were encountered: (a) the Jebsen test was insensitive to some important changes in hand function; (b) the test was sensitive to additional variables other than those directly related to hand function; (c) the test was affected by learning; (d) the tasks were not representative of actual ADL for tetraplegic patients; (e) the results were inconsistent since each task was only performed once; and (f) the test had inadequate instructions for application in tetraplegic patients.

Thirdly, available information on reproducibility, validity and sensitivity should be considered. Good reproducibility of test results is important, because otherwise an improvement in score cannot be solely attributed to the intervention, but can also be the result of variability in test scores. Validation of a test is important to determine whether a test actually measures what it claims to measure. High sensitivity of a test is important because the test has to be able to detect a change in function; otherwise it is useless as an instrument to evaluate an intervention.

A fourth consideration is the performability of a test by tetraplegic persons. For example, most tests require a sitting position during test performance, whereas during the initial assessment often only a lying position is possible. Another example is that some tests include items that are too difficult to perform by tetraplegic subjects. This problem was noticed both in general and disease specific tests. The modified Barthel Index (MBI)^{67,68} included two out of 15 items which could not be performed by tetraplegic persons, namely 'stair climbing' and 'walking'. Although the MBI is validated and seems sensitive enough to detect changes in time, the fact that some items cannot be performed by tetraplegics is a disadvantage. Incapacity of subjects to perform all the test items makes available results regarding reliability and validity less applicable. A disease-specific test in which this problem arose is the Thorson's functional test.⁶² The results of only three subjects were reported. The item 'volar grip of a bottle and drinking from it' was either not tested or could not be performed with or without the stimulatory device in five out of six measurements. The item does not seem to add much information, and therefore exclusion of the item should be considered.

Furthermore, additional characteristics of a test can play a role in selection of a test. For example, the amount of publications available, the date of publication, and the number of research groups using the test. Unfortunately, a large amount of experience is inaccessible to other investigators, because it remains unpublished.

If ADL performance is considered to be the outcome measure several ADL tests can be chosen. Only ADL tests which are specially designed for tetraplegics, or have been validated in tetraplegics should be used. Use

of the iv-BI, MBI, FIM, QIF and SCIM is preferred, because information is available on the psychometric properties of these tests when used in tetraplegics, in contrast to the RLAH Functional Activities Test, the COT and the VFM. As mentioned before, a disadvantage of the adapted versions of the Barthel Index is the inclusion of items which cannot be performed by tetraplegics. The FIM is useful when a general ADL scale is preferred. A disadvantage of the FIM over the QIF is that the QIF feeding scale may allow the detection of changes in function as individuals recover, that the FIM scale would miss.⁸¹ Yavuz *et al*⁸⁰ concluded that some additions to the FIM may be useful, especially in the feeding and dressing categories, and that a category of bed activities should be included, in order to improve sensitivity. A disadvantage of the FIM over the SCIM is that the FIM is less sensitive to functional changes in incomplete tetraplegics, except for the self-care subscore, in which the FIM and SCIM appear to have the same sensitivity.⁸⁵ Dependent on the composition of the intervention group the SCIM (aimed at all SCI persons) or the QIF (aimed at tetraplegics) is preferred.

Information on new tests or additional information on existing tests will become available in the coming years. A useful instrument to judge the quality of a new test is provided by Rudman and colleagues.⁸⁸ They gave an extended description of the criteria according to which a test can be evaluated. These criteria were put in a flow chart, which is called an instrument evaluation framework for selecting measurement instruments in hand therapy. Included in this evaluation tool were the following categories: clinical utility, standardisation, purpose, psychometric properties, and patient's perspective.

Some general considerations in selecting the appropriate evaluation test are discussed below.

Several functional tests use time as outcome measure. It is questionable that time is the most valid predictor of hand function.⁸⁹ Several arguments support this statement. Firstly, Smith²⁶ pointed out that although one hand may move in a slower, less coordinated manner than the other, the subject's ability to accomplish bilateral daily living skills within the normal time range is not necessarily hindered. Secondly, time seems to be less valid as an outcome measure due to the relation between speed and accuracy (the speed-accuracy trade-off).⁹⁰ An increase in speed is not necessarily the result of an improvement in function, but can also result from a decrease in accuracy with which the test is performed. And thirdly, the speed with which a task can be performed is probably less important to a patient than the ability to perform a task properly.

The functional tests discussed in the present review are meant to give an indication of upper extremity function. It should be kept in mind that, although such tests give some information about arm-hand function, the results cannot be generalised to arm-hand function during activities of daily living.

Jacobson-Sollerman and Sperling⁹¹ demonstrated this by testing 30 healthy subjects with the Ranchos Los Amigos (RLA) test, which is a modification of the hand function test developed by Carroll.³¹ The subjects were allowed a free choice of grip instead of those prescribed when the test is used for clinical purposes. The results show that healthy subjects performing the RLA test used a stereotyped grip pattern and several of the hand grips most frequently used in normal daily life (as determined in a standardised meal study) were rarely used. One reason for this seems to be the uniform nature of the test regarding the purpose of the handgrips, ie moving objects from and to various specified positions. The study confirmed previous reports that the action to be performed determines the choice of grip.

Validity of a test can be determined by comparing the test with a so-called 'gold standard'. When no gold standard is available, the alternative is to compare the test to another test with the same purpose. A disadvantage of this method is that although comparison makes clear that both tests do or do not measure the same aspect of function, no conclusions may be drawn regarding what is actually measured. For example, comparison of the Purdue Pegboard test to the Nine-Hole Peg test results in significant correlations.⁴⁷ This still does not justify the conclusion that these tests actually measure dexterity.

It is well known that the results of strength tests are influenced by factors like age and gender. In functional tests these influences can also be present. For example test results on the Smith hand function evaluation²⁶ are influenced by gender (female subjects worked faster than males, males demonstrated a stronger hand grip than females) and age (older persons worked slower than younger subjects). The Jebsen test revealed similar differences.⁴⁵ Additionally, normative values for tetraplegics can be influenced by disease specific factors, like differences in sensibility.⁷ Therefore, it seems important to include all these factors to facilitate a correct evaluation of the function of the hand.

Currently, much research has been performed with the purpose of improving arm-hand function in tetraplegic persons. Investigators do not always use the most appropriate tests to evaluate the effect of an intervention. The information as given in the present review can be used by investigators to select an evaluation tool(s) from the broad selection of evaluation tests that are available. In the future, more research should be performed with the purpose of improving tests or providing more information on psychometric properties of available tests when used in a tetraplegic population.

References

- 1 Hanson RW, Franklin MR. Sexual loss in relation to other functional losses for spinal cord injured males. *Arch Phys Med Rehabil* 1976; **57**: 291-293.

- 2 Deyo RA, Patrick DL. Barriers to the use of health status measures in clinical investigation, patient care, and policy-research. *Med Care* 1989; **27** (Suppl): S254–S268.
- 3 McPhee SD. Functional hand evaluations: a review. *Am J Occup Ther* 1987; **41**: 158–163.
- 4 Wade DT. Measuring arm impairment and disability after stroke. *Int Disabil Stud* 1989; **11**: 89–92.
- 5 Curtin M. Development of a tetraplegic hand assessment and splinting protocol. *Paraplegia* 1994; **32**: 159–169.
- 6 Krajnik SR, Bridle MJ. Hand splinting in quadriplegia: current practice. *Am J Occup Ther* 1994; **46**: 149–156.
- 7 Sollerman C, Ejeskar A. Sollerman hand function test. A standardised method and its use in tetraplegic patients. *Scand J Plast Reconstr Surg Hand Surg* 1995; **29**: 167–176.
- 8 Medical Research Council. Aids to the investigation of the peripheral nervous system. memorandum no. 45. Her Majesty's Stationery Office: London 1976.
- 9 Maynard Jr, FM, et al. International standards for neurological and functional classification of spinal cord injury. American spinal injury association. *Spinal Cord* 1997; **35**: 266–274.
- 10 Noreau L, Vachon J. Comparison of three methods to assess muscular strength in individuals with spinal cord injury. *Spinal Cord* 1998; **36**: 716–723.
- 11 Waters RL, Adkins RH, Yakura JS, Sie I. Motor and sensory recovery following complete tetraplegia. *Arch Phys Med Rehabil* 1993; **74**: 242–247.
- 12 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 Rehabil* 1992; **73**: 1063–1068.
- 13 Herbison GJ, Isaac Z, Cohen ME, Ditunno JF. Strength post-spinal cord injury: myometer vs manual muscle test. *Spinal Cord* 1996; **34**: 543–548.
- 14 May LA, Burnham RS, Steadward RD. Assessment of isokinetic and hand-held dynamometer measures of shoulder rotator strength among individuals with spinal cord injury. *Arch Phys Med Rehabil* 1997; **78**: 251–255.
- 15 Marciello MA et al. Wrist strength measured by myometry as an indicator of independence. *J Neurotrauma* 1995; **12**: 99–106.
- 16 Drolet M, Noreau L, Vachon J, Moffet H. Muscle strength changes as measured by dynamometry following functional rehabilitation in individuals with spinal cord injury. *Arch Phys Med Rehabil* 1999; **80**: 791–800.
- 17 Bohannon RW. Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20 to 79 years. *Arch Phys Med Rehabil* 1997; **78**: 26–32.
- 18 Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. *J Hand Surg [Am]* 1984; **9**: 222–226.
- 19 Mathiowetz V et al. Grip and pinch strength: normative data for adults. *Arch Phys Med Rehabil* 1985; **66**: 69–74.
- 20 Desrosiers J, Bravo G, Hebert R, Dutil E. Normative data for grip strength of elderly men and women. *Am J Occup Ther* 1995; **49**: 637–644.
- 21 Vanden Berghe A, Van Laere M, Hellings S, Vercauteren M. Reconstruction of the upper extremity in tetraplegia: functional assessment, surgical procedures and rehabilitation. *Paraplegia* 1991; **29**: 103–112.
- 22 Byers GJ, Goldstein BS, Sanders JE. An electromechanical testing device for assessment of hand motor function. *IEEE Trans Rehabil Eng* 1998; **6**: 88–94.
- 23 Memberg WD, Crago PE. Instrumented objects for quantitative evaluation of hand grasp. *J Rehabil Res Dev* 1997; **34**: 82–90.
- 24 de Castro MC, Cliquet Junior A. An artificial grasping evaluation system for the paralysed hand. *Med Biol Eng Comput* 2000; **38**: 275–280.
- 25 Bell E, Juerk K, Wilson T. Hand skill measurement: a gauge for treatment. *Am J Occup Ther* 1976; **30**: 80–86.
- 26 Smith HB. Smith hand function evaluation. *Am J Occup Ther* 1973; **27**: 244–251.
- 27 van der Ploeg RJ, Oosterhuis HJ. Physical examination: measurement of muscle strength. *Ned Tijdschr Geneesk* 2001; **145**: 19–23.
- 28 Keating JL, Matyas TA. The influence of subject and test design on dynamometric measurements of extremity muscles. *Phys Ther* 1996; **76**: 866–889.
- 29 Fleishman EA. Abilities and motor skill. In: Fleishman EA (ed). *The structure and measurement of physical fitness*. Prentice-Hall, Inc.: Englewood Cliffs, N.J. 1964, pp 23–24.
- 30 Gloss DS, Wardle MG. Use of the Minnesota Rate of Manipulation Test for disability evaluation. *Percept Mot Skills* 1982; **55**: 527–532.
- 31 Carroll D. A quantitative test of upper extremity function. *J Chron Dis* 1965; **18**: 479–491.
- 32 Lyle RC. A performance test for assessment of upper-limb-function in physical rehabilitation treatment and research. *Int J Rehabil Res* 1981; **4**: 483–492.
- 33 Carroll D. Hand function in quadriplegia. *Md State Med J* 1967; **16**: 107–108.
- 34 Tiffin J, Asher EJ. The Purdue Pegboard: Norms and studies of reliability and validity. *J Appl Psychol* 1948; **32**: 234–247.
- 35 Buddenberg LA, Davis C. Test–retest reliability of the Purdue Pegboard Test. *Am J Occup Ther* 2000; **54**: 555–558.
- 36 Desrosiers J, Herbert R, Bravo G, Dutil E. The Purdue Pegboard Test: normative data for people aged 60 and over. *Disability and Rehabilitation* 1995; **17**: 217–224.
- 37 Tiffin J. Purdue Pegboard Examiner Manual. Chicago: Science Research Associates; 1968.
- 38 Yeudall LT, Fromm D, Reddon JR, Stefanyk WO. Normative data stratified by age and sex for 12 neuropsychological tests. *J Clin Psychol* 1986; **42**: 918–946.
- 39 Mathiowetz V et al. The Purdue Pegboard: norms for 14- to 19-year-olds. *Am J Occup Ther* 1986; **40**: 174–179.
- 40 Jebsen RH et al. An objective and standardized test of hand function. *Arch Phys Med Rehabil* 1969; **50**: 311–319.
- 41 Spaulding SJ et al. Jebsen Hand Function Test: performance of the uninvolved hand in hemiplegia and of right-handed, right and left hemiplegic persons. *Arch Phys Med Rehabil* 1988; **69**: 419–422.
- 42 Colyer RA, Kappelman B. Flexor pollicis longus tenodesis in tetraplegia at the sixth cervical level. A prospective evaluation of functional gain. *J Bone Joint Surg Am* 1981; **63**: 376–379.
- 43 Mulcahey MJ, Smith BT, Betz RR, Weiss AA. Outcomes of tendon transfer surgery and occupational therapy in a child with tetraplegia secondary to spinal cord injury. *Am J Occup Ther* 1995; **49**: 607–617.
- 44 Noronha J, Bundy A, Groll J. The effect of positioning on the hand function of boys with cerebral palsy. *Am J Occup Ther* 1989; **43**: 507–512.

- 45 Agnew PJ, Maas F. Hand function related to age and sex. *Arch Phys Med Rehabil* 1982; **63**: 269–271.
- 46 Kellor M et al. Hand strength and dexterity. *Am J Occup Ther* 1971; **25**: 77–83.
- 47 Mathiowetz V, Weber K, Kashman N, Volland G. Adult norms for the Nine Hold Peg Test of finger dexterity. *Occ Th J Res* 1985; **5**: 24–38.
- 48 Holser P, Fuchs E. Box and Block test. In: Cromwell FS (ed). *Occupational therapists manual for basic skills assessment: primary prevocational evaluation*. Fair Oaks Printing Company: Pasadena, California 1960, pp 29–31.
- 49 Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block Test of manual dexterity. *Am J Occup Ther* 1985; **39**: 386–391.
- 50 Desrosiers J et al. Validation of the Box and Block Test as a measure of dexterity of elderly people: reliability, validity, and norms studies. *Arch Phys Med Rehabil* 1994; **75**: 751–755.
- 51 Kwakkel G et al. Intensity of leg and arm training after primary middle-cerebral-artery stroke: a randomised trial. *Lancet* 1999; **354**: 191–196.
- 52 Wagenaar RC et al. The functional recovery of stroke: a comparison between neuro-developmental treatment and the Brunnstrom method. *Scand J Rehabil Med* 1990; **22**: 1–8.
- 53 Van der Lee JH et al. The intra- and interrater reliability of the action research arm test: a practical test of upper extremity function in patients with stroke. *Arch Phys Med Rehabil* 2000; **82**: 14–19.
- 54 de Weerd WJG, Harrison MA. Measuring recovery of arm-hand function in stroke patients: a comparison of the Brunnstrom Fugl-Meyer-test and the Action-Research-Arm test. *Phys Canada* 1985; **37**: 65–70.
- 55 Broeks JG, Lankhorst GJ, Rumping K, Prevo AJ. The long-term outcome of arm function after stroke: results of a follow-up study. *Disabil Rehabil* 1999; **21**: 357–364.
- 56 Hsieh CL, Hsueh IP, Chiang FM, Lin PH. Inter-rater reliability and validity of the Action Research arm test in stroke patients. *Age and Ageing* 1998; **27**: 107–113.
- 57 Thrope G, Stroh K, Baco A. Standardized Object Test for quantitative assessment of hand grasp function using the C.W.R.U. upper extremity neuroprosthesis. In: Presperin JJ (ed). *Proceedings of the RESNA*; New Orleans 1989, pp 181–182.
- 58 Stroh K, Van Doren C, Thrope G, Wijman C. Common object test: a functional assessment for quadriplegic patients using an F.N.S. hand system. In: Presperin JJ (ed). *Proceedings of the RESNA*; New Orleans 1989, pp 387–388.
- 59 Stroh Wuolle K et al. Development of a quantitative Hand Grasp and Release Test for patients with tetraplegia using a hand neuroprosthesis. *J Hand Surg* 1994; **19A**: 209–218.
- 60 Smith BT, Mulcahey MJ, Betz RR. Quantitative comparison of grasp and release abilities with and without functional neuromuscular stimulation in adolescents with tetraplegia. *Paraplegia* 1996; **34**: 16–23.
- 61 Marino RJ, Shea JA, Stineman MG. The Capabilities of Upper Extremity instrument: reliability and validity of a measure of functional limitation in tetraplegia. *Arch Phys Med Rehabil* 1998; **79**: 1512–1521.
- 62 Thorsen R, Ferrarin M, Spadone R, Frigo C. Functional control of the hand in tetraplegics based on residual synergistic EMG activity. *Artif Organs* 1999; **23**: 470–473.
- 63 Post MW, van Asbeck FW, van Dijk AJ, Schrijvers AJ. Dutch interview version of the Barthel Index evaluated in patients with spinal cord injuries. *Ned Tijdschr Geneesk* 1995; **139**: 1376–1380.
- 64 Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. *Int Disabil Stud* 1998; **10**: 61–63.
- 65 Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Maryland State Med J* 1965; **14**: 61–65.
- 66 Granger CV, Albrecht GL, Hamilton BB. Outcome of comprehensive medical rehabilitation: measurement by PULSES profile and the Barthel Index. *Arch Phys Med Rehabil* 1979; **60**: 145–154.
- 67 Yarkony GM et al. Benefits of rehabilitation for traumatic spinal cord injury. Multivariate analysis in 711 patients. *Arch Neurol* 1987; **44**: 93–96.
- 68 Yarkony GM et al. Functional skills after spinal cord injury rehabilitation: three-year longitudinal follow-up. *Arch Phys Med Rehabil* 1988; **69**: 111–114.
- 69 Hamilton BB et al. A uniform national data system for medical rehabilitation. In: Fuhrer MJ (ed). *Rehabilitation outcomes. Analysis and measurement*. Paul H. Brookes Publishing Co.: Baltimore 1987, pp 137–147.
- 70 Hamilton BB, Laughlin JA, Granger CV, Kayton RM. Interrater Agreement of the Seven Level Functional Independence Measure (FIM). *Arch Phys Med Rehabil* 1991; **72**: 790.
- 71 Menter RR et al. Impairment, disability, handicap and medical expenses of persons aging with spinal cord injury. *Paraplegia* 1991; **29**: 613–619.
- 72 Muslumanoglu L et al. Motor, sensory and functional recovery in patients with spinal cord lesions. *Spinal Cord* 1997; **35**: 386–389.
- 73 Hall KM et al. Characteristics of the Functional Independence Measure in traumatic spinal cord injury. *Arch Phys Med Rehabil* 1999; **80**: 1471–1476.
- 74 Ota T et al. Functional assessment of patients with spinal cord injury: measured by the motor score and the Functional Independence Measure. *Spinal Cord* 1996; **34**: 531–535.
- 75 Warschausky S, Kay JB, Kewman DG. Hierarchical linear modelling of FIM instrument growth curve characteristics after spinal cord injury. *Arch Phys Med Rehabil* 2001; **82**: 329–334.
- 76 Rogers JC, Figone JJ. Traumatic quadriplegia: follow-up study of self-care skills. *Arch Phys Med Rehabil* 1980; **61**: 316–321.
- 77 Gresham GE et al. Quadriplegia Index of Function. *Arch Phys Med Rehabil* 1980; **61**: 493.
- 78 Labi MLC et al. Quadriplegia Index of Function: one-year follow-up. *Arch Phys Med Rehabil* 1981; **62**: 532–533.
- 79 Gresham GE et al. The Quadriplegia Index of Function (QIF): sensitivity and reliability demonstrated in a study of thirty quadriplegic patients. *Paraplegia* 1986; **24**: 38–44.
- 80 Yavuz N, Tezyurek M, Akyuz M. A comparison of two functional tests in quadriplegia: the quadriplegia index of function and the functional independence measure. *Spinal Cord* 1998; **36**: 832–837.
- 81 Marino RJ et al. Assessing selfcare status in quadriplegia: comparison of the quadriplegic index of function (QIF) and the functional independence measure (FIM). *Paraplegia* 1993; **31**: 225–233.

- 82 Marino RJ, Goin JE. Development of a short-form Quadriplegia Index of Function scale. *Spinal Cord* 1999; **37**: 289–296.
- 83 Mulcahey MJ *et al*. Implanted functional electrical stimulation hand system in adolescents with spinal injuries: an evaluation. *Arch Phys Med Rehabil* 1997; **78**: 597–607.
- 84 Catz A *et al*. SCIM – spinal cord independence measure: a new disability scale for patients with spinal cord lesions. *Spinal Cord* 1997; **35**: 850–856.
- 85 Catz A *et al*. The spinal cord independence measure (SCIM): Sensitivity to functional changes in subgroups of spinal cord lesion patients. *Spinal Cord* 2001; **39**: 97–100.
- 86 Catz A *et al*. The Catz-Itzkovich SCIM: a revised version of the spinal cord independence measure. *Disabil Rehabil* 2001; **23**: 263–268.
- 87 Taricco M *et al*. Functional status in patients with spinal cord injury: a new standardized measurement scale. Gruppo Interdisciplinare Valutazione Interventi Riabilitativi. *Arch Phys Med Rehabil* 2000; **81**: 1173–1180.
- 88 Rudman D, Hannah S. An instrument evaluation framework: description and application to assessments of hand function. *J Hand Ther* 1998; **11**: 266–277.
- 89 Bear-Lehman J, Abreu BC. Evaluating the hand: issues in reliability and validity. *Phys Ther* 1989; **69**: 1025–1033.
- 90 Schmidt RA. Methodology for studying motor behavior. In: *Motor control and learning. A behavioural emphasis*. 2nd edn. Human kinetics publishers, Inc.: Champaign, Illinois 1988, pp 65–66.
- 91 Jacobson-Sollerman C, Sperling L. Grip function of the healthy hand in a standardized hand function test. A study of the Rancho Los Amigos test. *Scand J Rehabil Med* 1977; **9**: 123–129.