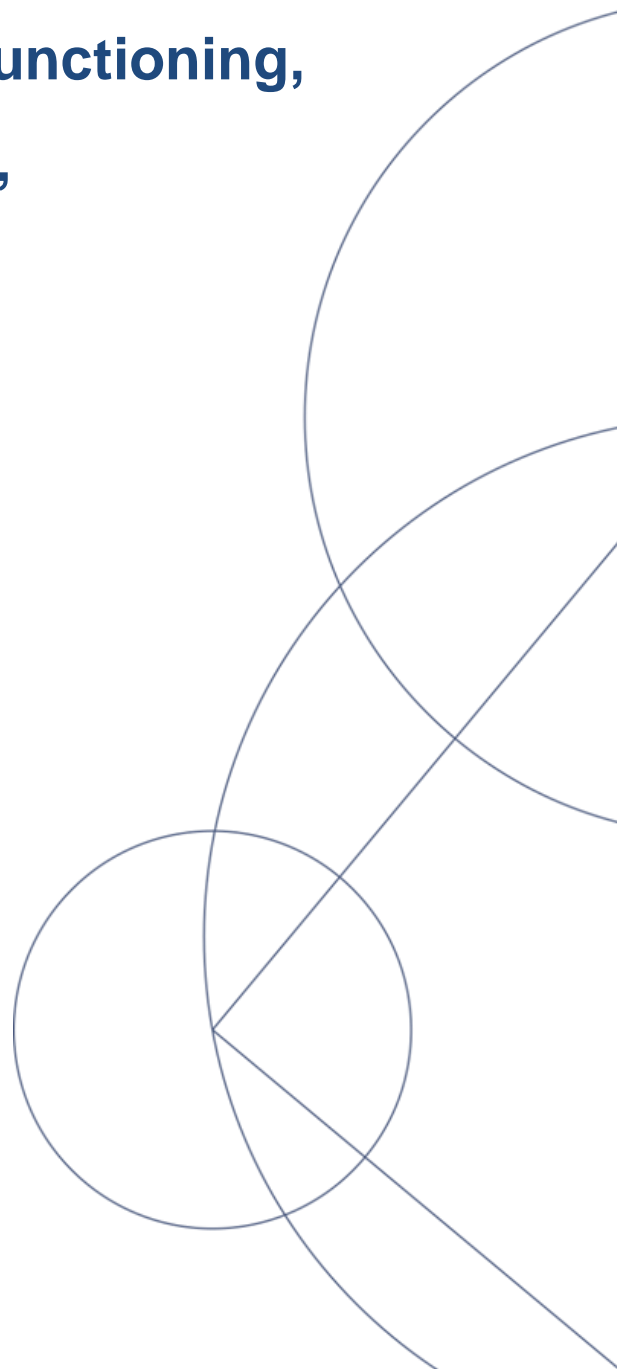




Ulla Werlauff

**Methods to assess physical functioning,
and their clinical applicability,
in patients with
spinal muscular atrophy and
congenital myopathy**

PhD thesis 2013



Methods to assess physical functioning, and their clinical applicability, in patients with spinal muscular atrophy and congenital myopathy

Ulla Werlauff, PT MSc

Academic Supervisor

Professor John Vissing, MD, DMSci, Neuromuscular Research Unit, Department of Neurology, Rigshospitalet, Copenhagen, Denmark.

Co-supervisor

Birgit Fynbo Steffensen, PT, PhD, the Danish National Rehabilitation Centre for Neuromuscular Diseases.

Assessment Comitte

Professor Fin Biering-Sørensen, MD DMSci Rigshospitalet, Copenhagen, Denmark.

Michelle Eagle, PT, PhD, Newcastle University, United Kingdom

Professor Mar Tulinius, MD, PhD, Göteborg University, Department of Pediatrics, Sweden

CONTENTS

PREFACE AND ACKNOWLEDGEMENT	4
ABBREVIATIONS	6
LIST OF PAPERS	7
SUMMARY ENGLISH.....	8
DANSK RESUME	9
INTRODUCTION	11
Assessment at impairment level	13
Assessment at activity level	14
Assessments at participation level	15
THE TWO DIAGNOSES	16
Congenital myopathy	16
Spinal muscular atrophy	17
AIMS	19
Specific aims.....	19
PATIENTS AND METHODS.....	21
Study I.....	21
Study II	22
Study III.....	22
Study IV	22
Ethics	23
Assessments at impairment level	24
Antropometrics (study I).....	24
Measurement of dynamic muscle strength (study I, II, III).....	24
Measurement of isometric muscle strength (study I, II)	25
Measurement of range of motion (study I, II)	25
Measurement of respiratory capacity (study I, III).....	26
Assessments at activity level	26
Brooke upper limb scale (Studies I, II, III)	26
Egen Klassifikation (studies I, II, III)	27
Hammersmith functional motor scale (study I).....	27
Motor Function Measure (study II).....	28
Assessments at participation level	28
Fatigue Severity Scale (study IV)	28
Visual Analog Scale (study IV)	29
Focus-groups (study IV)	29

Data analyses / Statistics.....	29
RESULTS.....	31
Assessments at impairment level.....	31
Dynamic muscle strength / Manual muscle test (study I, II, III).....	31
Isometric muscle strength / Hand held Dynamometry (studies I, II).....	33
Range of motion in upper limbs (studies I, II).....	33
Respiratory capacity (studies I and III).....	34
Assessments at activity level.....	35
Brooke upper limb scale (studies I, II, III).....	35
EK scale (studies I, II, III).....	36
Hammersmith functional motor scale (study I).....	37
Motor Function Measure (study II).....	38
Assessments at participation level.....	39
Fatigue Severity Scale.....	39
Visual analogue scale.....	39
Reliability.....	39
Content validity.....	39
Construct validity.....	41
MAIN RESULTS.....	42
Study I.....	42
Study II.....	42
Study III.....	42
Study IV.....	43
DISCUSSION.....	45
Generic versus disease specific scales.....	45
Scale level.....	46
Assessments at impairment level.....	47
Manual muscle test.....	47
Hand held dynamometry.....	48
Range of motion.....	48
Respiratory capacity.....	49
Assessments at activity level.....	49
Brooke upper limb scale.....	50
Egen Klassifikation scale.....	50
Hammersmith functional motor scale.....	50
Motor Function Measure.....	51
Assessments at participation level.....	52
Fatigue Severity Scale.....	52
Visual analogue scale.....	52
CONCLUSION.....	55
REFERENCES.....	56
APPENDIX – THE FUNCTIONAL SCALES, PAPERS.....	69

PREFACE AND ACKNOWLEDGEMENT

Three years have passed since I became PhD student at the University of Copenhagen. Looking back, time has flown but has at the same time been very intense, filled with data collections, courses, reflections, preoccupations, writings and discussions – in other words learning; learning about methodology and methods and especially about measurements. “To measure is to know¹”, so what cannot be measured cannot be understood. In the studies encompassed by this PhD thesis, some of the measurements used in the clinical assessment of patients with neuromuscular diseases have been analyzed; hopefully the results will contribute to the knowledge and understanding of the two diagnoses that have been the interest of the studies.

I am very grateful for the support and the interest shown by colleagues, family and friends throughout the study but I wish to express my special thanks to:

All participants in the four studies. Thank you for your willingness to participate and share your knowledge. You have made this study possible.

John Vissing, MD Professor and my supervisor, thank you for your supervision, for scientific guidance, for valuable and instructive comments in my draft manuscripts and for language correction. I have learned a lot from your expertise.

Birgit Steffensen, PT PhD, my co-supervisor, co-author, mentor and colleague at the Danish National Rehabilitation Centre for Neuromuscular Diseases, and an authority within the field of neuromuscular diseases. Thank you for sharing your vast knowledge, for your support, encouragement, enthusiasm and constructive criticism and for always being there when I am in a puzzle and need to discuss something.

Jes Rahbek, MD, director at the Danish National Rehabilitation Centre for Neuromuscular Diseases, and my boss. Thank you for gently pushing me into the direction of research and for your support when I got there.

¹ Lord Kelvin 1824-1907

Ann-Lisbeth Højberg, OT MSi, co-author in the paper on fatigue, colleague and friend. It is a great pleasure and always inspiring to work with you. Thank you for caring and for our enriching discussions, also on subjects that not relate to research.

Susanne Berthelsen, Ida Fløytrup, Bente Kristensen and Birgit Werge, PTs', colleagues and co-authors in the cross-sectional study on SMA. Thank you for your collaboration and for collecting data for the longitudinal study.

All my colleagues at the Danish National Rehabilitation Centre for Neuromuscular Diseases. After periods "out of house" I have always enjoyed being back in the office with good colleagues, good talks and good laughs. It means a lot.

Susanne Rosthøj, statistician PhD, for helping me with the statistical calculations in the longitudinal paper.

Rasmus Firla-Holme, co-author in the fatigue paper, for working out the statistics in the analysis of the psychometric properties of the Fatigue Severity Scale.

Gert, my husband, thank you for your support and understanding when days were long and sometimes gloomy, and **Christine and Kasper**, thank you for your encouragement and firm belief that I could do it.

This study was supported by the Danish National Rehabilitation Centre for Neuromuscular Diseases and Danske Fysioterapeuters Forskningsfond.

ABBREVIATIONS

CIS	Checklist individual strength
CM	Congenital myopathy
DMD	Duchenne muscular dystrophy
EK	Egen Klassifikation
EK2	Egen Klassifikation – extended version
ENMC	European Neuromuscular Centre
FSS	Fatigue Severity Scale
FVC	Forced vital capacity
HFMS	Hammersmith functional motor scale
HHD	Hand Held Dynamometer
ICF	International Classification of Functioning, Disability and Health
MFM	Motor Function Measure
MFM D3	Motor Function Measure - Distal dimension
MMO	Maximal mouth opening
MMT	Manual muscle test
MRC	Medical Research Council
MS	Multiple Sclerosis
NMD	Neuromuscular Diseases
PCA	Principal Component analysis
PRO	Patient reported outcome measures
ROM	Range of motion
SLE	Systemic lupus erythematosus
SMA	Spinal muscular atrophy
SMN	Survival motor neuron
VAS	Visual Analog Scale

LIST OF PAPERS

List of papers

- I. Werlauff U, Steffensen BF, Bertelsen S, Fløytrup I, Kristensen B, Werge B: Physical characteristics and applicability of standard assessment methods in a total population of spinal muscular atrophy type II patients. *Neuromuscul Disord* 2010; 1: 34–43.
- II. Werlauff U, Steffensen BF. The applicability of four clinical standard methods to evaluate arm and hand function in all stages of Spinal Muscular Atrophy type II. *In review*
- III. Werlauff U, Vissing J, Steffensen BF. Change in muscle strength over time in spinal muscular atrophy types II and III. A long-term follow-up study. *Neuromuscul Disord* 2012; 12: 1069-74.
- IV. Werlauff U, Højberg A, Firla-Holme R, Steffensen BF, Vissing J. Fatigue in Patients with Spinal Muscular Atrophy type II and Congenital Myopathies; Evaluation of the Fatigue Severity Scale. *In review*.

SUMMARY ENGLISH

Neuromuscular disorders encompass a variety of sub groups in which impaired muscle strength is the principal symptom, that in most cases are caused by mutations in genes that affect the neuromuscular unit. Despite very different phenotypes, a common feature to all neuromuscular disorders is the impact on muscle strength, which influences all domains of function as defined in the International Classification of Function, disability and health. When the consequences of a neuromuscular disorder is evaluated, it is thus necessary to describe both capacity – defined as the impact on body functions - and capability – the impact on activity and participation. This puts demands on the assessment methods used for evaluation, which alone or in combination should be able to provide a holistic picture of the patient.

The two disorders of interest in this thesis are spinal muscular atrophy (SMA) and congenital myopathy (CM). SMA represents a group of very weak patients in whom the natural course of disease hasn't been well described due to lack of responsiveness in the methods used for evaluation of impairment and activity. Congenital myopathy is an umbrella term that covers a range of subtypes with similarities and differences but with experienced fatigue as a general clinical symptom, which seems to impact activity and participation, although this has never been investigated systematically.

The four studies encompassed by this thesis investigated how the characteristic features in the two disorders can be evaluated. The results from studies I-III contribute to the knowledge on the natural history of SMA, and to the knowledge about which clinical assessment methods are the most applicable to evaluate impairment and activity in these weak patients. The results suggest that upper limbs - where muscle strength and functions are best preserved in SMA - should be the area of focus if changes over time or results of interventions should be evaluated. Study IV concerns the impact of fatigue as perceived by patients with SMA and CM. The hypothesis of fatigue being a problem in patients with CM was confirmed, and the applicability of an existing instrument to evaluate fatigue in the two disorders was illustrated. Experienced fatigue is easy to assess and provide important information on impact of function from the patients' perspective.

DANSK RESUME

Neuromuskulære sygdomme eller muskelsvind er en betegnelse for en række sygdomsgrupper, hvor hovedsymptomet er nedsat muskelkraft forårsaget af genetiske mutationer, der påvirker funktionen af den motoriske enhed. Der er stor variation i graden af nedsat muskelkraft afhængig af den enkelte muskelsvindstype, men det er en fælles betingelse, at den nedsatte muskelkraft har konsekvenser for funktionsevnen, som denne er beskrevet i den internationale klassifikation af funktionsevne, funktionsevnenedsættelse og helbredstilstand.

Når man skal undersøge konsekvenserne af at have en neuromuskulær sygdom, er det således nødvendigt at beskrive såvel påvirkning af kropsfunktion som påvirkning af funktion på aktivitets- og deltagelsesniveauet. Det stiller krav til de metoder, man anvender til undersøgelsen, idet metoderne hver for sig eller i fællesskab skal kunne give et fuldt og dækkende billede af patienten. I dette studie har der været fokus på to forskellige neuromuskulære sygdomme, spinal muskel atrofi (SMA) og kongenit myopati (CM). Diagnosen SMA omfatter en gruppe patienter med en væsentlig påvirkning af muskelkraften. Sygdomsforløbet er ikke vel beskrevet, hvilket til dels skyldes at de undersøgelsesmetoder, der har været anvendt ikke har været tilstrækkelig følsomme til at beskrive funktionsevnen og ændringer i denne. Kongenit myopati er en overordnet diagnose, der rummer en række undertyper med såvel ligheder som forskelle. Det er imidlertid generelt, at patienter med CM oplever, at træthed er en faktor, der influerer på de daglige funktioner og har stor betydning i hverdagen. Dette er aldrig blevet undersøgt.

De fire studier, der er omfattet af denne afhandling, undersøger hvordan de karakteristiske træk ved de to former for neuromuskulære sygdomme kan evalueres og beskrives. Resultaterne fra studierne I-III bidrager med viden om naturhistorien ved SMA og hvilke kliniske undersøgelsesmetoder, der kan anvendes til at undersøge og beskrive funktion hos personer med meget lidt muskelkraft.

Resultaterne viser, at muskelkraft og funktion er bedst og længst bevaret i armene, og at målinger derfor skal koncentreres om armene, hvis der skal opfanges forandringer over tid eller som resultat af en behandling.

Studie IV undersøger, hvordan patienter med SMA og CM oplever træthed og trætheds indflydelse på hverdagsfunktioner. Resultaterne bekræfter hypotesen om, at patienter med CM oplever træthed som et problem, og studiet viser, at en eksisterende metode til at undersøge træthed kan anvendes ved de to typer af neuromuskulære sygdomme. Oplevet træthed er nemt at undersøge og giver vigtig information om påvirkning af funktionsevnen set fra patienternes synspunkt.

INTRODUCTION

Neuromuscular diseases are a group of hereditary disorders that affect different parts of the neuromuscular unit, which include the spinal cord motor neurons, the peripheral nerve, the neuromuscular junction and the muscles. The degree of physical impairment is dependent on the specific type of disorder. Since no cure has been found, treatment aims at minimizing the consequences of the primary muscle weakness and to preserve functional ability at all stages of the disease. The overall aim is that the patients can reach or sustain their optimal level of independence and function [United Nations 1994]. In order to develop and evaluate interventions that make this possible, it is essential to have a documented knowledge of the natural history of the specific disease and a thorough understanding on how functional ability is influenced by the disease and related factors.

The overall aims of the clinical examination are to determine the extent of physical impairment, to monitor the course of the disease and to evaluate the impact on the patients' daily life. This examination is the patient's first contact with the hospital and is central for the supportive examinations that are initiated to reach a diagnosis. The assessment methods that are used to evaluate the neuromuscular patients therefore need to be adequately informative to quantify the characteristics of the individual disease, and adequately sensitive to discriminate among patients and register any gain or loss of function. To be relevant and meaningful for the patient, it is also important that the assessments reflect the patient's function in the perspective of daily living [Fowler 1982]. To reflect the impact of a disease is thus not only a matter of measuring capacity – defined as biological function – but also to measure capability, the ability to perform and engage in activities [Whitbeck 1978].

This understanding corresponds to the World Health Organizations definition of "Health", as "a state of complete physical, mental and social well-being, and not merely as the absence of disease" [WHO 1948]. Consequently, it is not sufficient to measure impact related to body function when a specific disease is evaluated and when rehabilitation is planned; impacts on activity and participation must also be measured. The international classification of Functioning, Disability and Health (ICF) [WHO 2001] is used as framework to cover these aspects. The two key concepts in ICF are "disability" and "rehabilitation", defined as:

Disability

- an umbrella term covering the following three areas : Impairment, activity limitations and participation restrictions where impairment is a problem in body function e.g. reduced muscle strength, activity limitations are difficulties in performing activities e.g. walking or eating and participating restrictions are problems in life situations e.g. accessibility. All three areas of disability must be assessed to plan rehabilitation.

Rehabilitation

- a set of measures that assist individuals who experience, or are likely to experience, disability to achieve and maintain optimal functioning in interaction with their environment.

Function

Motor performance is a dynamic interaction between numbers of interdependent elements, which can be understood by analyzing these elements. As an example, an individual's ability to eat on his own is dependent on muscle strength, joint motion, coordination, and whether he has respiratory capacity and energy to perform the function.

Patients with neuromuscular diseases and impaired muscle strength make use of compensatory strategies. It is essential to understand and to evaluate these strategies since they are necessary to maintain a function; furthermore the degree of compensatory strategies often express the current stage of the disease. If focus is solely on the level of impairment, there is a risk of losing information on a gradual change in motor performance; vice versa would an instrument solely focusing on activity not explain the cause of the change in motor performance. As an example, an examination of muscle strength in the forearm will illustrate the capacity to flex the elbow, but not the patient's capability to bend the elbow and lift the hand to the mouth; the latter being a composite function of muscle strength and often performed by compensatory movements. In this way function may be defined as *“the ability to interact with ones environment in a way that permits the person to achieve competence in the tasks of daily living. Underlying this is an expression of the person's physical competence to control the physical components of muscle strength and range of motion. In the very weak patients the interdependence of these parameters means that small or subtle changes in any component have a disproportionate effect on function. It is therefore necessary to have scales*

which alone or in combination provide a holistic picture of the patient and his performance”
[Sylvia Hyde, personal communication 2001].

In the clinical setting, as well as in research, it is important to decide the relevant outcome - what should be measured? Is it capacity, capability or both? In rehabilitation research there are often several outcomes of interest; the primary outcome is usually at the level of activities, and the outcomes at impairment level act as a gold standard and help to interpret other findings. Consequently, the applicability of the assessment methods depends on their ability to measure the outcomes [Wade 2003].

Assessment at impairment level

Methods used to assess level of impairment are related to body function and physical capacity. One of the primary characteristics in neuromuscular disorders is impaired muscle strength and it is essential to be able to quantify muscle strength in order to diagnose the patient, to pick up changes over time and to evaluate treatment. Muscle strength can be measured by manual or by quantitative techniques and as dynamic or isometric muscle strength. The manual muscle test (MMT) was developed by Lovett (1916) and has since become a widely used method. It has been slightly moderated over the years, and is typically scored using the 6-point Medical Research Council (MRC) scale with 0 representing “no muscle function” and 5 representing “normal” muscle strength [Medical Research Council 1943]. Scores from the tested muscles are often summed into a composite score to express overall muscle strength. The test is useful as a screening tool to evaluate muscle strength and to plan for rehabilitation; it is easy to perform and does not require any form of equipment. The scores from “0” – “3”, representing the most decreased muscle strength, are well defined whereas scores above “3” are based on a subjective decision made by the evaluator. As a consequence, inter and intra-rater agreement is better when testing weaker muscles compared to stronger muscles [Florence 1992, Mahoney 2009]; agreement is further improved when the test is performed by trained evaluators [Florence 1984, Escolar 2001]. In general, studies have provided evidence for good reliability and validity in the use of MMT [Cutberth 2007], but the increasing demands for objectivity based on interval scaling in measuring muscle strength have limited the use of MMT in clinical trials in neuromuscular diseases.

Quantitative muscle strength is considered to be an “objective” measurement of muscle strength. The method is highly correlated with MMT [Saraniti, 1980, Aitkens 1989, Goonetilleke 1994] and is as such often considered as a gold standard in muscle strength

assessments. The test score is recorded by means of a mechanical device and can be compared to reference values obtained from persons with normal strength. Several forms of quantitative measurements exist. The most applicable version in a clinical setting is the portable hand held dynamometer (HHD) that measures isometric muscle strength in a standard position. The isometric measurements can be performed as “make-test” in which the evaluator holds the dynamometer stationary and the patient exerts a pressure, or as “break-test” in which the evaluator pushes the dynamometer and the patient holds against the push. The break-test triggers a slightly higher force than the make-test [Bohannon 1988, Stratford 1994], but the make-test is easier to control. In general, tests of reliability have shown excellent agreement among evaluators, but also that reliability is dependent on the evaluator’s experience and can be seriously affected if the standardization of the test and the stabilisation of the dynamometer is not met [van der Ploeg 1991, Merlini 2002, Mahoney 2009].

Assessment at activity level

Methods used to assess the level of activity, aim at measuring motor performance. In children, the methods are often based on the knowledge about when motor milestones are reached in the healthy child, and aim at identifying delays in these milestones [Henderson and Sugden 1992, Folio and Fewell 2000, Nelson 2006]. Since the primary symptom in neuromuscular disorders is reduced muscle strength, the scales must be able to reflect how this affects motor performance. In view of the fact that longevity of patients with neuromuscular diseases has increased, it is necessary that the scales can be used on patients of all ages and with a wide range of functional abilities – also in patients with very limited muscle strength.

Several specific performance-based measurements have been developed with the purpose of evaluating motor function in neuromuscular disorders. One of the earliest functional scales is the Vignos scale [Vignos 1963] that evaluates ambulation in boys with Duchenne muscular dystrophy (DMD). Additional scales were developed in the 1980’s e.g. Brooke upper limb scale that classifies upper limb function in neuromuscular disorders [Brooke 1981] and Hammersmith motor ability scale that evaluates motor ability in DMD [Scott 1982]. In the last decade, further functional scales have been developed and are now part of the assessment battery in neuromuscular disorders. The Egen Klassifikation scale (EK) [Steffensen 2001] assesses motor function related to daily activities in non-ambulatory patients with DMD or SMA, The Hammersmith functional motor scale

(HFMS) for SMA assesses motor function in non-ambulant patients with SMA [Main 2003] and The Motor Function Measure (MFM) assesses motor function in patients with all neuromuscular disorders [Berard 2005]. Extended and/or moderated versions of the HFMS and the EK scale have been developed [Krosschell 2006, Steffensen 2008], and there is an ongoing effort to develop clinical methods that can evaluate activities relevant for the patients [Mazzone 2011] and has psychometric properties that can qualify the scale to act as an outcome measure in clinical trials.

Assessments at participation level

A neuromuscular disease affects more than physical function. Daily life, family life, working life and social relations are also influenced. The degree and the importance of these impacts cannot be measured objectively; so to obtain information and improve the understanding on these impacts, the patients must be asked on their opinion and experiences. Patient reported outcome measures (PRO) are recommended, already widely used and often a request in clinical trials [European Medicines Agency 2005, US Food and Drug Administration 2009, Rothman 2009, Patrick 2011]. The patient perspective is increasingly included in research and patients and patient organizations are now involved in the making of Standard of Care programs and clinical trials protocols [Wang 2010, McCormack 2013]. Patient involvement in scientific research priorities for neuromuscular disorders indicates that research must be balanced between fundamental research in health and factors that influence health (such as fatigue) and research on quality of life [Nierse 2013].

Muscle fatigue and low endurance are commonly described in neuromuscular disorders, but research on fatigue and the impact on daily function have in general received little attention [Féasson 2006, Lou 2010]. Since severe fatigue can impact on all domains of function, this must be addressed in order to determine supportive measures [Wokke 2007, de Vries 2010]. Experienced fatigue can be qualified by (PRO) instruments and/or by qualitative interviews [Pettersson 2009]. There is a variety of scales that assess fatigue and the impact of fatigue, but since none have been developed specifically to quantify fatigue in neuromuscular disorders, a number of scales have been used, some generic and some developed for use in other disorders. Two of the scales most often used to quantify whether fatigue is a symptom in neuromuscular disorders are the Fatigue Severity Scale (FSS) [Lou 2010, Laberge 2005, Hagemans 2007] and the Checklist individual strength (CIS) [Kalkmann 2005, Schillings 2007]. The FSS seeks to evaluate fatigue at one level, namely impact of fatigue on daily functioning, whereas the CIS evaluates fatigue in four

levels: subjective fatigue, concentration, motivation and physical activity. Both scales have been recommended for use in future studies on fatigue in neuromuscular disorders [ENMC 2011] to further evaluate their applicability in these disorders.

Very few studies have used qualitative interviews to address the experiences of living with a neuromuscular disease in terms of consequences for activity and participation, but two studies [Boström 2004, Heatwole 2012] emphasize that fatigue restricts activities of daily life. Qualitative interviews can be one of the first steps to develop a PRO instrument, but can also act as method to validate an existing instrument to assure that the concept of interest captures the concept from the patients' perspective [Patrick 2011].

THE TWO DIAGNOSES

Congenital myopathy

Congenital myopathies (CM) are a group of neuromuscular diseases in which symptoms typically appear at birth or in infancy. The incidence and prevalence of CM is unknown. The diseases are caused by mutations in genes that encode proteins involved in the contractile function of muscle fibers and are grouped in four “morphological subclasses” based on features seen on muscle biopsy (myopathies with protein accumulations – e.g. rods, cores, central nuclei and fiber size variations) [North 2008]. Genes have been identified within each of the four subgroups, but finding the right genetic diagnosis has been complicated by the fact that one gene mutation can cause different histological features and different gene mutations can cause the same histological feature.

The clinical presentation of the CMs is at a continuum of severity varying from a slight to severe impairment of physical function. Despite different genetic backgrounds, the phenotypes show great overlap among the different CMs although some clinical characteristics are specific for the individual type and may lead to genetic testing. As an example, does the finding of fiber size variation combined with early scoliosis and impaired respiratory function indicate a SEPN1 mutation.

The classical clinical features in CM are impaired muscle strength and delayed motor development with onset at birth or from early childhood. Mimics are often impaired resulting in ophthalmoparesis, ptosis and/ or bulbar involvements [Ryan 2001, Jungbluth 2005], but there is a great variability in the degree of these symptoms and some types of CM do not affect facial muscles. In general, motor development progresses after the initial period of life, but then stabilizes. Some patients may, however experience a continuous loss of physical function or worsening of symptoms later in the disease course. The degree and range of clinical symptoms as well as the natural course of the diseases is still not well understood due to lack of data on the natural history of the diseases.

Despite being a heterogeneous group of disorders, some common impairments influence daily life in patients with congenital myopathies. Some frequent complaints are fatigue and low endurance [Wang 2012], but no systematic studies on these complaints have been conducted in this group of disorders.

Spinal muscular atrophy

Spinal muscular atrophy (SMA) is an inherited neuromuscular disease characterized by degeneration of the spinal cord motor neurons, which is caused by a mutation in the SMN1 gene. The incidence of SMA is approximately 1:10.000 [Thieme 1993, Arkblad 2009]. A copy gene (SMN2) may influence the severity of the disease [Feldkotter 2002, Wirth 2006], but SMN2 copy number cannot be used to predict the clinical phenotype [Wang 2007], since this can vary considerably within the same copy number. The clinical spectrum of SMA ranges from severe hypotonia and weakness to mild weakness and the disease is classified into three types according to age of onset and achievement of motor milestones [Munsat 1992].

SMA I – also called Werdnig-Hoffman disease - is the most severe type with onset before the child is 6 month old. Motor function is severely impaired, the child never sits and because of bulbar dysfunctions and pulmonary complications most children die during the first two years of life.

SMA II is the intermediate type with onset before the child is 18 month old. The child will achieve the ability to sit independently, but not to stand or walk unaided. Some children lose their ability to sit independently at an early age whereas others maintain the ability into adulthood.

SMA III – also called Kugelberg-Welander disease - is the milder type with onset from the 18th month. The child achieves the ability to walk independently, but in some children this ability is lost in early childhood whereas others maintain ambulation into adulthood.

The phenotypic spectrum of SMA represents a continuum, and there is a wide range of functional abilities within each SMA subtype, consequently some borderline type I/II and type II/III do exist. A further sub-classification has been suggested [Dubowitz 1995, Zerres 1995, Russman 1996], but not decided; however, the classification of an adult type, SMA IV, with onset in the second or third decade is now often used [Wang 2007].

The distribution of muscle weakness in II and III is well described [Carter 1995, Kroksmark 2001], but the natural history of SMA II and III has not been studied systematically. There is a general agreement that patients lose functional abilities over time due to the changes in growth and secondary complications as scoliosis [Dubowitz 1995, Zerres 1995, Russmann 1996]. Despite electrophysiological studies have indicated an aged-related loss of innervation in SMA [Swoboda 2005], opinions differ on whether muscle strength also deteriorates. Some studies have indicated loss of muscle strength over time as measured by manual muscle test [Carter 1995, Steffensen 2002, Deymeer 2008], whereas other studies could not demonstrate deterioration in muscle strength over time, when muscle strength was assessed by quantitative methods [Iannaccone 2000, Kauffman 2011].

This diversity could be due to the fact that various outcome measures and various time spans have been used to study the course of the disease in patients that cover a wide field of disability from hardly any measurable muscle strength to nearly normal muscle strength.

AIMS

The overall aims of this thesis were to evaluate the applicability of standard assessment methods and their ability to evaluate and reflect physical characteristics in patients with spinal muscular atrophy. A second aim was to evaluate the prevalence and impact of fatigue in spinal muscular atrophy and congenital myopathy.

Specific aims

- To describe muscle strength, functional capability, contractures and Forced Vital Capacity in a total population of SMA II patients (study I)
- To evaluate the applicability of standard assessment methods and their ability to detect variations in muscle strength and functional ability among these individuals (studies I, II).
- To evaluate decline in muscle strength over time in SMA II and III (study III).
- To investigate whether fatigue is a common feature in patients with SMA II and CM, and whether the Fatigue Severity Scale is an appropriate instrument to identify and evaluate this (study IV).

PATIENTS AND METHODS

Study I and II are cross-sectional studies. Study III is a retrospective study on longitudinal data.

Study IV is a mixed- method study on validity of the Fatigue Severity Scale. Validity was evaluated by statistic analyses and focus group interviews.

Study I describes the applicability of standard assessment methods to characterize physical function primarily at impairment level. Study II describes the ability of clinical methods to reflect capacity and capability at impairment and activity level. Study III describes the sensitivity of clinical methods to register loss of muscle strength at impairment and activity level. Study IV evaluates the ability of a questionnaire to reflect impact of fatigue on levels of activity and participation.

Criteria for inclusion in study I, II, and III were that the patients;

- had a genetically confirmed diagnosis of spinal muscular atrophy
- had a clinical diagnosis of SMA type II (study I, II and III) and SMA III (study III) based on the established diagnostic criteria [Munsat 1997]
- were \geq five years of age at time of first examination

Criteria for inclusion in study IV were that patients:

- were diagnosed with a congenital myopathy based on muscle biopsy and/or molecular diagnosis
- were diagnosed with SMA II based on clinical and genetic findings
- were \geq eighteen years of age

Study I

The total Danish population of 67 patients with SMA II, registered with the Danish National Rehabilitation Centre for Neuromuscular Diseases in august 2007, were invited to participate in this study. Data were obtained from 54 participants (21 females, 33 males) aged 5 – 70 years. All

patients invited, fulfilled the diagnostic criteria for SMA II and all patients had a homozygous deletion of the SMN1 gene. Among the 54 patients, three patients had two copies of the SMN2 gene, 36 patients had three copies and 14 patients had four copies.

All assessments were undertaken at the Rehabilitation Centre and were conducted by six experienced physiotherapists, who worked in pairs. Before the examinations, patients had filled in registration forms with information on spinal surgery, respiratory and nutritional problems and respiratory and nutritional aids.

Study II

The total population of Danish patients with a clinical and genetically confirmed diagnosis of SMA II and registered with the Danish National Rehabilitation Centre for Neuromuscular diseases in September 2010 (n= 65), were invited to participate in this study. Data were obtained from 52 participants (8 - 73 years). The majority of patients were assessed at the Centre, but a few patients were assessed at home. All assessments were undertaken with the patients in their wheelchair and by the same physiotherapist.

Study III

Data from 23 patients with SMA II and seven patients with SMA III that had participated in 2 - 6 different studies on muscle strength and motor function during the last twenty years were analyzed. Median follow-up was 17 years (12-20). Median number of assessment was 4 (2-6). All assessments had been performed by the same four experienced physiotherapists. Measurements that had been used in all studies were used for analyses.

To assess whether the baseline level of muscle strength at entry had an influence on potential progression, SMA II patients were divided in two groups according to upper limb function at entry.

Study IV

Twenty-nine patients with SMA II and 71 patients with CM \geq 18 years filled in the Fatigue Severity Scale (FSS). Data on SMA II patients were obtained from study II; data on CM patients were obtained from a study on CM conducted at the Neuromuscular Research Unit, Rigshospitalet, and

the Danish National Rehabilitation Centre for Neuromuscular diseases in 2010. In both studies, patients had filled in the FSS questionnaire at time of the examination. The validity and the reliability of the FSS were examined by a combination of quantitative and qualitative methods; as part of the validation process, twelve patients with CM reported their experience on fatigue in two focus-group interviews.

Tabel I. Patient characteristics in the four studies. *Age at last assessment

	Data collection	Number	Mean age (range)	Female /male	Type of data	ICF domains of function
Study I	2007				Cross sectional	Impairment
SMA II		54	23 (5 – 70)	21/33		
Study II	2010-2011				Cross sectional	Impairment Activity
SMA II		52	26 (8 – 73)	22/30		
Study III	1991 - 2011				Longitudinal	Impairment Activity
SMA II		23	38 (22 – 73)*	9/14		
SMA III		7	43 (28 – 62)*	5/2		
Study IV	2010 - 2011				Questionnaire	Participation
SMA II		29	31 (19 – 55)	10/19		
CM		71	34 (18 – 73)	36/35	Focus groups	

Ethics

The studies were approved by the ethics committee of the Capital Region and the regional ethics committees. Written information was sent to the patients together with the invitation to the studies, and verbal information was given to the patients at the examinations/interviews. Patients signed an informed consent. If the patient was < 18 years, parents signed the consent.

Assessments at impairment level

Gender and ages were recorded and reported in all studies. Influence of gender was calculated in study IV. In study I, the patients' age was used as cut points to describe characteristics in different age groups and to calculate differences among young and old patients.

Antropometrics (study I)

Since all patients in study I were non-ambulant, height in meters was measured as full length of arm span between finger tips by a flexible measuring tape [Hepper 1965, Miller 1992]. Weight was recorded in kilos using a scale for lifts.

Measurement of dynamic muscle strength (study I, II, III)

Muscle strength was tested by manual muscle testing (MMT) based on the Medical Research Council (MRC) method [1943]. MRC scores 0-5 were modified to a 0-10 score to make the scale more sensitive according to the modification of Brooke et al. [1981]. Manual muscle testing of 38 muscle groups of the whole body was measured in study I, and based on our findings in this study, we decided to focus on muscle strength of the upper limbs in study II, in which nine muscle groups were measured. In study III, we calculated the change over time in seven muscle groups of the upper limbs in patients with SMA II and III, and for SMA III also in six muscle groups of the lower limb. All muscle tests were performed bilaterally. MRC score was calculated as percentage of maximal possible score by the fraction used by Scott et al [1982]: $MRC\% = \frac{\text{sum of graded scores}}{\text{number of muscle tested} \times 10} \times 100$.

The reliability of the manual muscle test increases when performed by a clinically experienced evaluator and when followed by a standardized protocol [Cuthberth 2007, Escolar 2001]. Inter-rater reliability is also improved when testing weak muscles and when tested by a limited number of experienced evaluators [Kleyweg 1991, Florence 1992, Mahony 2009].

These criteria were all met in our studies: Patients with SMA II have very limited muscle strength ranging from 1 to 3 on the MRC scale. All assessments were performed by one of four trained and clinically experienced evaluator and training sessions were arranged over the years to ensure consistency and agreement among evaluators. At the last training session, agreement among four

evaluators was estimated. Scores were very consistent among evaluators with a variance of 0-8% (median 4%) in MRC score.

Measurement of isometric muscle strength (study I, II)

Hand held dynamometer (HHD Citec™, C.I.T.Technics BV, Groningen, the Netherlands) was used to measure force in Newton (N) as the maximal voluntary contraction. The method was chosen to reflect muscle strength at interval level. Four muscle groups were measured: elbow flexors, elbow extensors and finger flexors/full fist grip (study I), and finger flexors and thumb muscles/lateral pinch grip (study II). The measurements were performed bilaterally as “make-test” with the patient and dynamometer application in standardized positions according to the Citec™ manual. Each measurement was repeated three times and the best score was recorded. According to the manufacturer’s direction, the grip score was multiplied with two. Reference values for the Citec™ dynamometer have been established [Beenakker et al 2001] as well as for other HHD’s [Bäckman 1989, 1995], but since our aim was to evaluate the applicability of the HHD method to reflect capacity in patients with SMA II, we did not compare the obtained values with reference values. The Citec™ dynamometer used in our clinic and in studies I and II is one of many hand-held dynamometers. The reliability of quantitative muscle test measured by HHD is in general considered to be high [Mafi 2012], also in neuromuscular disorders [Iannaccone 2000, Merlini 2002, Mahoney 2009], but reliability is influenced by the muscles measured, the strength of the muscle, the placing of the applicator and by the evaluator’s own capacity [Wadsworth 1987, Bohannon 1988, Mahoney 2009].

In study II, the internal consistency of the HHD - measured by Chronbach alpha – was 0.997 for full fist grip and 0.991 for lateral pinch grip.

Measurement of range of motion (study I, II)

Passive range of motion (ROM) in the upper limbs was measured by a standard goniometre according to the methodology established by the American Academy of Orthopaedic Surgeons [1965] in shoulder flexion (only study I), elbow extension/supination,/pronation – hand flexion/extension/ulnar /radial deviation and finger extension. Contractures were calculated as the difference between recorded ROM and the normal ROM [Brooke 1981]. Hypermobility was

defined as the ability to bend joint passively beyond normal range of motion. In study I, maximal mouth opening (MMO) was measured in millimetres as the distance between lower and upper incisors.

The reliability of measurement of range of motion is high when the measurements follow a standardized protocol and when performed by the same investigator; studies indicate that measurements of upper limbs are more reliable than measurements of lower limbs [Boone 1978, Gajdosik 1987].

Measurement of respiratory capacity (study I, III)

Respiratory capacity was measured as forced vital capacity (FVC) by means of a calibrated spirometer (Medikro Spiro2000). The FVC expresses the volume of air forcibly exhaled in one breath, following a maximal inhalation. Patients were measured in sitting and supine position since patients with SMA II in general have an increase in FVC in supine position compared to sitting position [Lyager 1995]. Each measurement was performed three times; best value was recorded and expressed as FVC% according to the reference value for the individual patient. In study III, change of FVC% over time was calculated as the difference between FVC% at first and last assessment. FVC is very reproducible and is widely used to reflect respiratory capacity in neuromuscular disorders [Samaha 1994, Wang 2007].

Assessments at activity level

Brooke upper limb scale (Studies I, II, III)

Brooke upper limb scale [Brooke et al. 1981] was used to evaluate and classify upper limb function. The patients' ability to move their arms independently were categorized into six levels; level 1 represents highest level of motor function and level 6 lowest level of motor function. In study I, the scale was used to illustrate the difference in upper limb function between younger patients \leq twenty years old and patients \geq twenty-one years. In study II, patients were classified according to Brooke upper limb scale to illustrate the wide range of upper limb function among patients with SMA II, and the classification was used to evaluate whether other clinical methods could reflect the various levels of upper limb function. In study III, patients were divided into two superior Brooke group's

(Brooke levels, 1, 2, 3 and Brooke levels 4, 5, 6) to assess whether the basic level of muscle strength had an influence on potential progression.

The Brooke upper limb scale is based on the ability to raise arm against gravity. The tasks at levels 3 and 4 can be performed by compensatory movements e.g. support from armrest. To improve reliability we defined tasks at these levels without use of armrest. The reliability of Brooke score in a SMA II population was estimated as agreement among four evaluators; there was total agreement among evaluators.

Egen Klassifikation (studies I, II, III)

Egen Klassifikation (EK scale) [Steffensen et al 2001, 2002] evaluates the non-ambulant patients' overall physical function. The scale was originally constructed with ten items, each representing daily activities relevant for the patient, and an extended version with seven added items has recently been developed (EK2) [Steffensen 2008]. The scale is administered to the patients as a combination of interview on daily performance and a visual examination of the tasks that can be observed at the assessment. Each item has four categories, scaled from 0 to 3, with 0 representing highest level of function and 3 lowest level of function. The EK sum is calculated as the sum of scores of all items. The EK scale (10 items) was administered to all patients in study I and the EK2 scale (17 items) was administered to all patients in study II. Five of 17 items evaluates activities in upper limbs, and in study II the sum of these five items was calculated as "EK2 upper limb module" (illustrated in the table section in study II). In study III, the differences between EK sum at first and last assessment were calculated.

Reliability has been established for the EK scale [Steffensen 2002] and the EK2 scale [Steffensen 2008]

Hammersmith functional motor scale (study I)

The Hammersmith functional motor scale (HFMS) for non-ambulant children with SMA was developed by Main et al [2003]. The scale has 20 items based on normal motor milestones and is scored in lying, sitting and standing position. Each item is scored from 0 to 2 based on the performance on the individual item. Maximum score is 40 corresponding to highest level of

function, minimum score is 0. The scale has been modified (MHFMS-SMA) with a slight change and reordering of some of the items [Krosschell 2006]. We used the original scale in our study. The reliability of the HFMS has been established in a multinational study [Mercuri et al 2006] with children from 2-12 years.

Motor Function Measure (study II)

The Motor Function Measure (MFM) scale was developed for patients with neuromuscular diseases to assess motor function across the range of mobility and the type of neuromuscular disorder. The scale has 32 items in three domains of function: standing and transfer, axial/proximal dimension and distal dimension. Each item is scored from 0 to 3, with 0 representing lowest level of function and 3 highest level of function; the MFM score is calculated as a percentage of highest possible score for each dimension and/ or for all dimensions. The distal dimension, MFM D3, was used in study II to evaluate upper limb function. This dimension has seven items, of which six items measure motor function in forearm and hand and one item evaluates the ability to dorsi-flex the foot. We chose to calculate the MFM D3 score with and without this item, the latter as the “MFM D3 upper limb” score (illustrated in the table section in study II).

The reliability of the MFM scale has been established for patients with neuromuscular disorders from 6-62 years [Berard 2005].

Assessments at participation level

Fatigue Severity Scale (study IV)

The Fatigue Severity Scale (FSS) was developed to assess the self-reported impact of fatigue on daily functioning in Multiple Sclerosis (MS) and systemic lupus erythematosus (SLE) [Krupp 1989]. The scale has nine items, each of which is rated on a 7-point Likert scale ranging from “1 = strongly disagree” to “7 = strongly agree.” The FSS score is calculated as the mean of all item scores. Based on results from normal populations and disease populations a FSS score ≥ 4 indicates that fatigue is a problem in daily life and a score of ≥ 5 indicates severe fatigue.

The reliability of the FSS has been established in normal populations [Lerdahl 2005, Valko 2008] and various disease populations [Krupp 1989, Kleinman 2000, Hagemans 2007], but has not previously been established in neuromuscular populations.

Visual Analog Scale (study IV)

The Visual Analog Scale (VAS) is used to measure a variety of subjective phenomena, among them fatigue. The score is rated on a 100 mm horizontal line with descriptions at each end, and represents the patient's perception on the phenomenon. The reliability of a VAS scale to measure fatigue severity has been presented in healthy individuals and patients with sleep-disorder [Lee 1991] and in patients with stroke [Tseng 2010].

Focus-groups (study IV)

Focus-groups were used to assess the content validity of the FSS and its relevance to measure impact of fatigue in patients with CM. A purposive sampling was performed to identify patients with FSS score ≥ 4 ; among these patients, 16 were randomly selected and invited to participate in one of two focus-group interviews. Interviews were based on an interview-guide with six main themes. For each theme, a set of 2-6 sub-questions were constructed, in case the group needed inspiration for the discussions. At the end of the discussion, the participants could add further comments on a piece of paper or write information they wouldn't bring up in the group. After a short break, the FSS was handed to each of the participants, who were then asked to comment on whether fatigue according to their perception was contained in the FSS, and if not – which dimensions of fatigue was not captured in the scale. The focus-group interviews were recorded on tape.

Data analyses / Statistics

Statistical analyses was conducted by means of Statistical Package for the Social Sciences (SPSS 16.0) in study I, SAS 9.2 software package (SAS Institute Inc) in studies II and III, and Stata version 11.2 in study IV. Significance levels were set at $p < 0.05$, using two-tailed testing. Non-parametric statistics were used when criteria for normal distribution were not met, and when the sample size was small. For calculation of unpaired differences between groups, Mann-Whitney U test was used in studies I and II and the parametric analogue, student t-test, in study IV. To calculate paired differences within groups, Wilcoxon signed rank test was used in studies I and III. Kruskal Wallis' test was used to calculate differences among groups in study II. Linear regression analyses were used to test longitudinal data in study III. Spearman's rank order correlation coefficient (r_s)

was used in studies I, II and IV. In study IV, principal components factor analysis (PCA) was used to test uni-dimensionality and Chronbach's alpha (α) coefficient was used to illustrate internal consistency.

Focus-group data were transcribed and analyzed by means of a direct content analyses [Krueger 2008, Hsieh 2005] and thematic prevalence. Validity was assessed as the relation between the themes emerging from the interviews and the nine FSS items.

Table II. Statistical methods used in the studies

	Study I	Study II	Study III	Study IV
Descriptive statistics				
Mean/SD	x			x
Median/Range	x	x	x	x
Analytic statistics				
Kruskal Wallis		x		
Wilcoxon signed rank test	x		x	
Mann-Whitney U test	x	x		
Student t-test				x
Linear regression			x	
Spearman's rank correlation	x	x		x
Principal component analysis				x
Cronbach alpha coefficient				x
Statistical software				
SPSS 16.0	x			
SAS 9.2		x	x	
Stata version 11.2				x

RESULTS

Assessments at impairment level

Dynamic muscle strength / Manual muscle test (study I, II, III)

All patients in studies I, II and III were measured by manual muscle test. In study I, the method confirmed a characteristic pattern of muscle involvement in spinal muscular atrophy type II: Upper limbs were relatively stronger than lower limbs and distal muscle groups were stronger than proximal muscle groups. Adductor muscles were the strongest muscles in shoulders and hips, elbow flexors were stronger than elbow extensors. In the lower limbs, knee flexors were weaker than knee extensors. The strongest muscles were found in the forearm (elbow flexors, wrist flexors/extensors and finger flexors/thumb muscles); muscles in fingers were the best preserved muscles, since all of the weakest patients had some muscle strength left in their thumbs and/or finger flexors. In general, a MRC% score of the upper limbs reflected a larger variation in capacity among patients compared to a MRC% score based on a full muscle test (fig. 1).

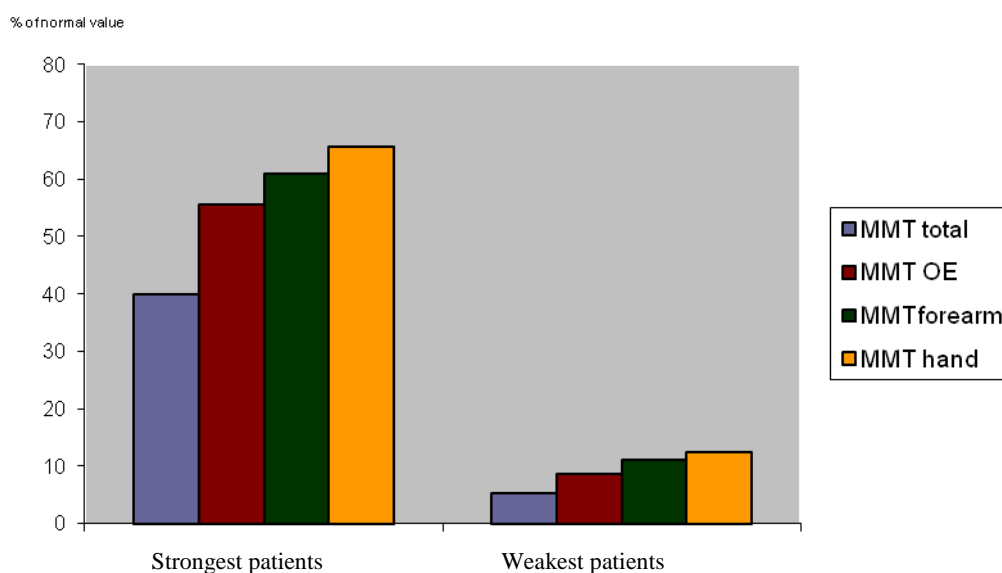


Figure 1. Mean MRC% score of the three strongest and three weakest patients in study I, recorded as MMT total (38 muscle groups), MMT OE (20 muscle groups) MMT forearm (12 muscle groups) MMT hand (8 muscle groups). MRC% score based on a reduced muscle test reflected a larger capacity than a total muscle test in both stronger and weaker patients.

In study II, MRC% score based on a reduced muscle test of the upper limb could discriminate among patients at all Brooke levels, also among the weakest patients at Brooke level 5 and 6 ($p = 0.002$). In the two cross-sectional studies (I and II) younger patients had more muscle strength than older patients ($p < 0.001$). This indicated a loss of muscle strength over time, a finding that was confirmed in study III, where analysis over time showed a loss of muscle strength in the upper limbs in patients with SMA II ($p < 0.0001$) and III ($p < 0.02$), respectively (figure 2). In patients with SMA II, the slope of the regression line was -0.22 units per year ($p < 0.03$). The rate of decline in muscle strength seemed a little steeper for stronger patients, but no significant difference was found. When repetitive MRC% scores of upper limbs, forearms and hands were compared, MRC% scores of forearm and hand seemed to be more sensitive to change over time (fig. 3).

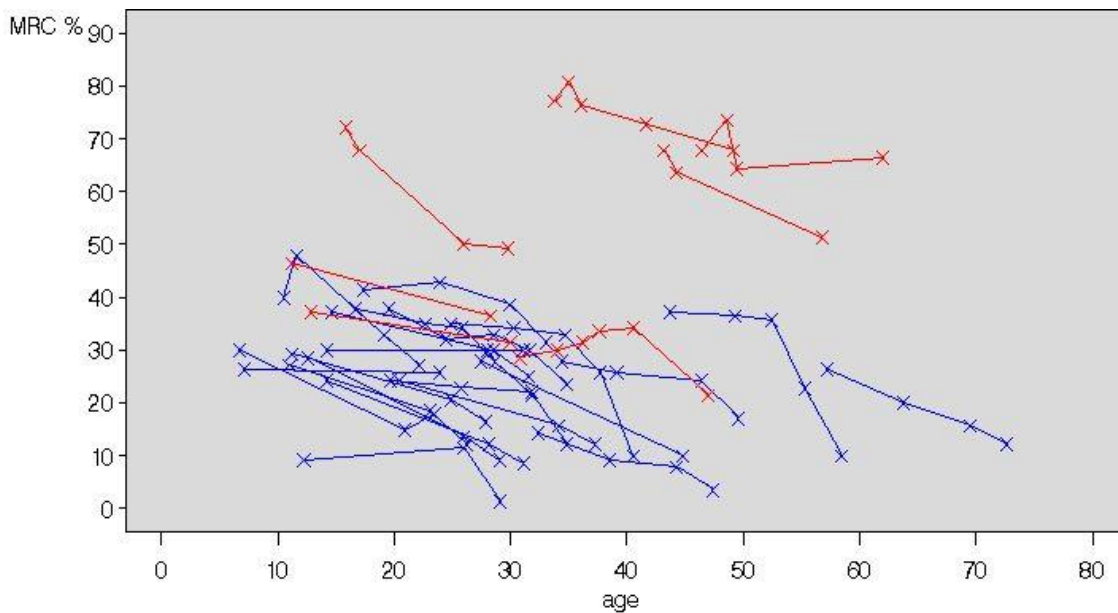


Figure 2. Decline in muscle strength over time. Patients with SMA II = blue lines, patients with SMA III = red lines.

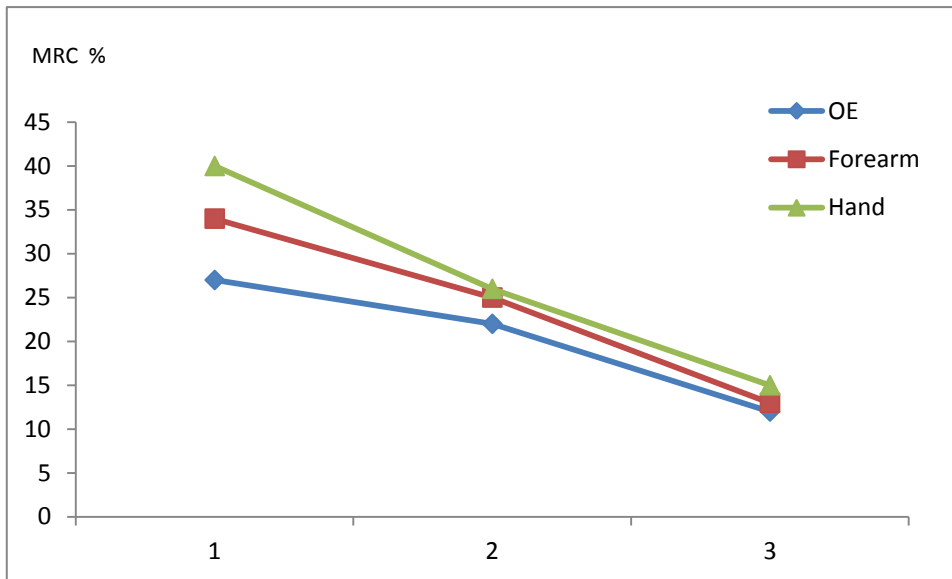


Figure 3. MRC% scores over time in 21 patients with SMA II. The colored lines represent median scores in 14/10/6 muscle groups of the upper limbs. Numbers on the X-axis refers to 1, 2, and 3 assessment. There was an interval of at least 4 years between each of the illustrated time intervals.

Isometric muscle strength / Hand held Dynamometry (studies I, II)

Not all patients could be measured by the Citec™ HHD since 18-40% of the patients measured in studies I and II could not overcome the dynamometer's activation threshold in any of the tests performed. Elbow flexion was scored by 82% of the patients in study I and was not performed in study II, where patients were measured sitting in their wheelchair and therefore could not obtain the standard position for the test. Full fist grip were scored by 67% of the patients in study I and 43% of the patients in study II. Lateral pinch grip was scored by 60% of the patients in study II. Neither of the two hand tests could reflect the capacity of all individuals nor could they discriminate among patients across the range of upper limb function, as classified by Brooke upper limb scale; full fist grip could, however, discriminate between patients at Brooke levels 3 - 4 and 4 - 5, respectively. Younger patients had higher scores than older patients in both studies.

Range of motion in upper limbs (studies I, II)

All patients - regardless of age or level of upper limb function - were hypermobile in fingers and in wrist ulnar flexion. More than 50% of the patients were hypermobile in elbow pronation. Limited ROM in shoulders was common in older patients. In general, older patients had more contractures

than younger patients, and asymmetry in contractures was more prominent ($p = 0.024$). Limited ROM in elbow extension/ supination, wrist extension/radial flexion was common in patients at Brooke levels 3-6. Contractures tended to be more prominent in the stronger arm and increased in general from Brooke levels 2–5; however, a significant difference in sum of contractures was only found between patients at Brooke level 3 and the adjoining level 4 ($p = 0.011$)

Limited mouth opening was found in 96% of the patients aged > 20 years.

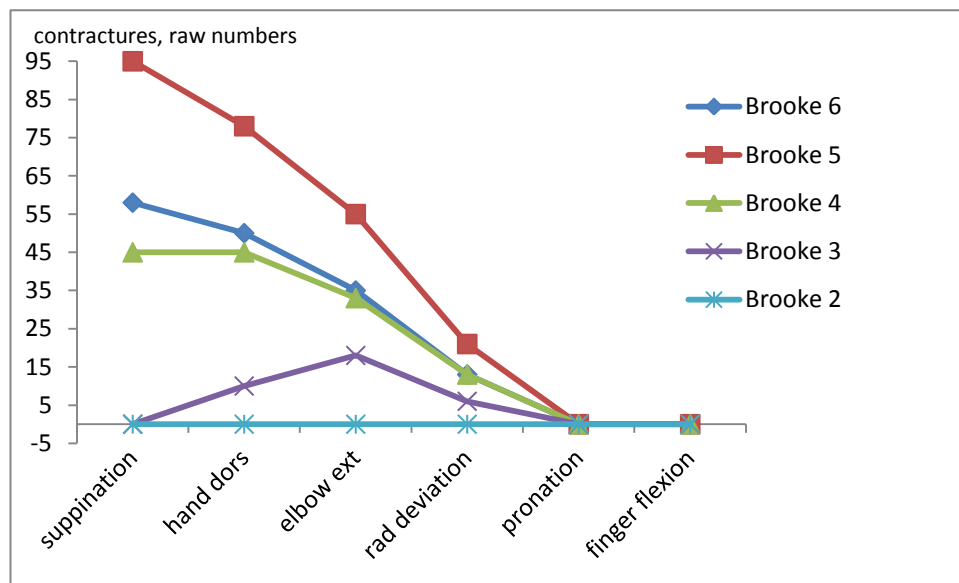


Figure 4. Joint motion in 52 patients (study II) illustrated as median contractures at five Brooke levels. No contractures were found in pronation and finger flexion despite the level of Brooke upper limb function. Contractures seemed to follow the level of function: the lower level the more contractures; however patients at Brooke level 5 had more contractures than patients at Brooke level 6. (Hypermobility is not illustrated).

Respiratory capacity (studies I and III)

In study I, forced vital capacity was measured in 42/54 patients. The remaining twelve patients were either unable to be without their ventilator or could only be measured in one of the standard positions (supine and sitting). 53% of the patients measured had the largest FVC in supine compared to the sitting position. 55% of the patients used assisted ventilation; 20% had invasive ventilation by tracheostomy and used the ventilator around the clock, 35% of the patients had non-invasive ventilation during night time. FVC% was highly correlated with EK sum score (-0.754) and moderately correlated with MRC% score (0.655) and functional tests (HFMS 0.677, Brooke

upper limb scale -0.627). There was no difference in FVC% mean value between younger patients and older patients with SMA II.

In study III, FVC% was evaluated over a period of 15 years (13-16) in ten patients with SMA II and five patients SMA III. Two of the SMA II patients were ventilated via tracheotomy at the last assessment. Consequently, FVC% from the previous assessment (two years before) was used for calculations in these two patients. There was no difference between first and last assessments for patients with SMA II ($p = 0.184$) nor for patients with SMA III ($p = 0.188$).

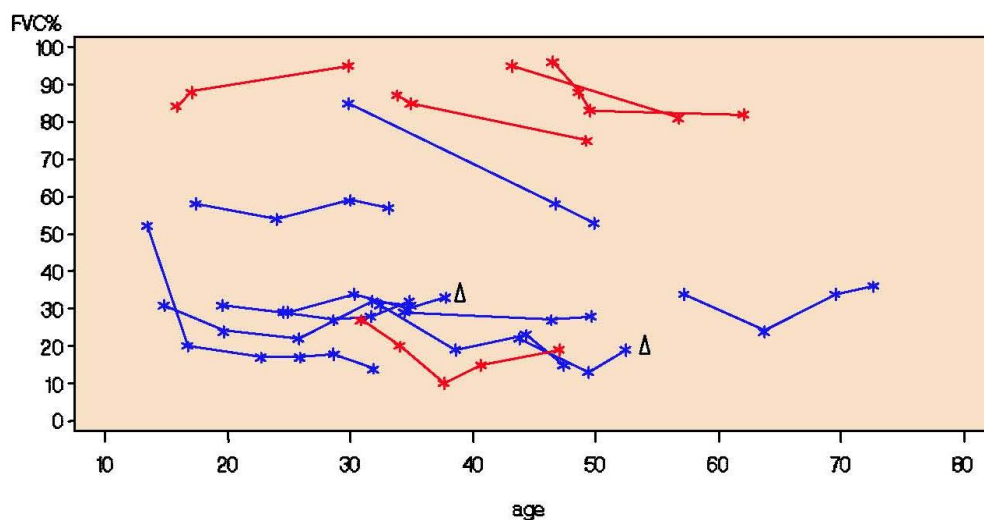


Figure 5 FVC% over time in SMA II patients (n = 10/blue) and SMA III patients (n = 5/red). Δ = patients with tracheostomy at last assessment. Change over time ($p = 0.184$, $p = 0.188$)

Assessments at activity level

Brooke upper limb scale (studies I, II, III)

Brooke score correlated highly with a total MRC% score (-0.885), and MRC% score of the upper limbs (0.887). Correlation with age was moderate (0.452); nevertheless younger patients had a higher level of upper limb function (lower Brooke score) compared to older patients ($p = 0.003$). Brooke score in individual patients changed significantly over time ($p < 0.0001$). Approximately 20% of the patients in studies I and II were categorized at Brooke level 6 corresponding to no useful function of hands. This indicates a floor effect in patients with SMA II, which was supported by the fact that several of these patients possessed upper limb capabilities, which could be measured by

other clinical methods (figure 6). 75% of the patients could drive their wheelchair by an adapted joystick, 50% of the patients could send text messages or use remote control (EK2), and 50% of the patients could slide their finger from one square to another (MFM).

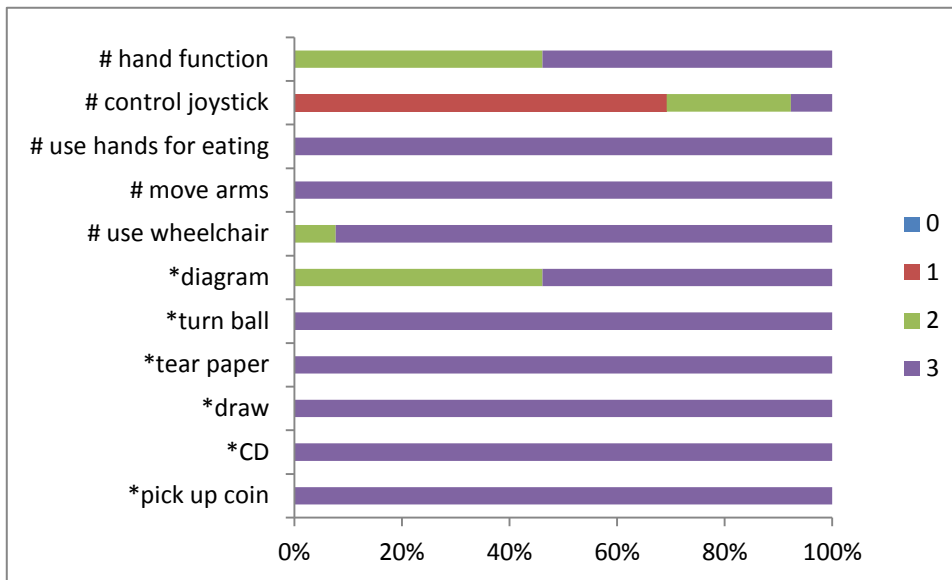


Figure 6. Upper limb capabilities in patients at Brooke level 6 (no useful function of hand) as measured by EK2 upper limb (#) and MFM upper limb (*). Each item was scored from 0 -3 with 3 representing minimal function. In three EK2 items and one MFM item, patients had capabilities that could be measured (Note that scores for MFM are reversed in this figure with “3” representing minimal function).

EK scale (studies I, II, III)

The EK sum score correlated highly with at total MRC% score (-0.845) and a MRC% score of the upper limbs (-0.877) (study I). EK2 sum score and MRC% score of the upper limbs correlated also highly (-0.917), and correlation was even better between EK2 upper limb score and MRC% score of the upper limbs (-0.958) (study II). Correlation between EK sum score and age was low (0.364), a finding that was confirmed in study II (0.393) for both EK and EK2 sum scores; correlation between age and EK2 upper limb score was slightly higher (0.513). Still, younger patients had a higher level of physical capabilities compared to older patients, and in study III, a significant change over time in EK sum score was shown ($p < 0.0001$).

Minimum and maximum EK sum scores were not in use in either of the EK scales, thus even the weakest patients have capabilities that could be reflected by the EK scales. Degree of difficulty for the individual item on the EK2 upper limb is illustrated in figure 7.

The EK2 sum score could discriminate among patients with very different upper limb capabilities, except the weakest patients, the EK2 upper limb score could discriminate among all patients (p values ranging from 0.001 – 0.005).

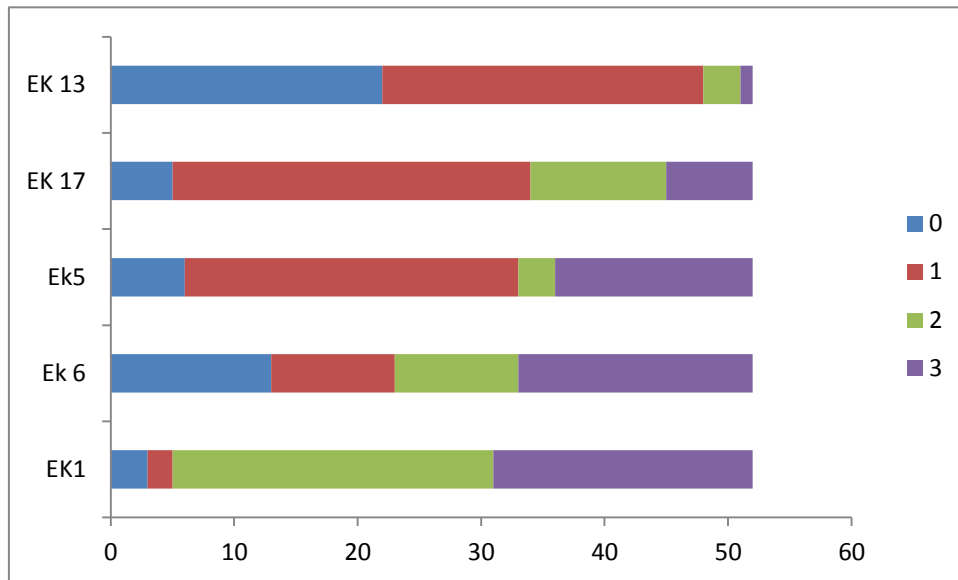


Figure 7. "EK2 upper limb"; five items – each scored from 0 -3 with 3 representing minimal function. Rank of difficulties for the individual items with the easiest item on top. The lower score the higher function. Item 13 (ability to control joystick) was the easiest item since 48/52 patients scored 0 or 1. Item 1 was the most difficult item; only 5/52 patients scored 0 or 1.

Hammersmith functional motor scale (study I)

The HFMS and the MRC% correlated well (0.735), although 44% of the patients scored 0 on each of the 20 HFMS items. Correlation with age was low (-0.453). Still, younger patients performed better than older patients (p = 0.004). Median HFMS score for 21/54 patients ≤ 16 years old was 4 (0-17). The median score improved slightly if patients were ≤ 12 years (17 patients /median score 5), but still 5/17 patients couldn't score on the scale indicating that the HFMS has a considerable floor effect in patients with SMA II.

Motor Function Measure (study II)

The MFM dimension 3 was highly correlated with MRC% score (0.925), the correlation improved when the item measuring foot motion was omitted (0.937). Correlation between age and MFM D3 score was low (-0.472), but younger patients ≤ 20 years had a significantly higher score than older patients ($p = 0.02$).

Minimum respective maximum MFM D3 scores were used in 3/52 patients. When the score was calculated for the six upper limb items only, minimum score was used in 7/52 patients, maximum score in five patients. Degree of difficulty for the individual item on the MFM upper limb is illustrated in figure 8.

The MFM D3 score could discriminate among all patients across the range of upper limb function as measured by Brooke upper limb scale (p values ranging from 0.001 – 0.032). The ability to discriminate among the strongest patients was lost in “the MFM upper limb score”.

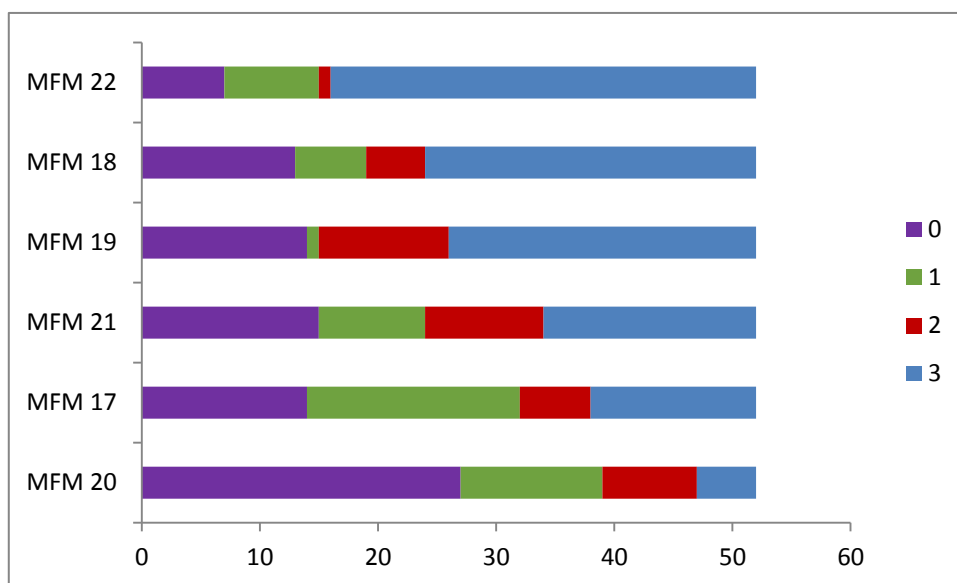


Figure 8. MFM upper limb; six items – each scored from 0-3 with 3 representing maximal function. Rank of difficulties for the individual items with the easiest item on top. The higher score the higher function. Item 22 (move fingers) was the easiest item since 37/52 patients scored 2 or 3. Item 20 (tear paper) was the most difficult item with 39/52 patients who scored 1 or 0.

Assessments at participation level

Fatigue Severity Scale

The prevalence of fatigue as measured by the FSS was high in patients with CM and low in patients with SMA II. The cut off FSS score ≥ 4 - indicating that fatigue is a problem in daily life - was scored by 76% of patients with CM, but only by 10% of the patients with SMA II. A FSS score ≥ 5 - indicating severe fatigue - was scored by 52% of patients with CM. There was no difference in FSS score between non-ambulant and ambulant patients with CM ($p = 0.21$), but women had higher FSS scores than men ($p = 0.04$). There were no gender differences in the SMA II group.

Visual analogue scale

The intensity of fatigue rated by the VAS scale was high in patients with CM and low in patients with SMA II. Median score for the CM group was 6.3 and 1.6 for the SMA II group. Ambulant patients with CM had a higher VAS score compared to non-ambulant patients ($p = 0.04$). There was no significant gender difference in either of the two groups.

Reliability

Reliability was assessed as agreement between two assessments. Correlations between test and retest of the FSS were 0.72 based on answers from ten patients with CM and 0.98 based on answers from six patients with SMA II. There were no significant differences in scores between first and second assessment in either of the groups. Out of a total of 144 pairs of ratings included in the test-retest analyses, 44% were identical and 28% were within ± 1 level of disagreement.

Correlation between test and retest of the VAS was very low (0.12) in six patients with CM and very high (0.99) in six patients with SMA II. In the CM group, the correlation was influenced by one patient who had diminished his score by 50% in the retest and one patient who had tripled his score in the retest.

Content validity

Each patient with CM responded to all nine FSS items, and the range of scores from 1-7 was used in all items. Among SMA II patients, 50% did not respond to item 2, and the range of scores from 1

- 7 was used in only three items (2, 3 and 4). Distributions of FSS scores for the two disease groups are illustrated in figure 9 a. and 9 b.

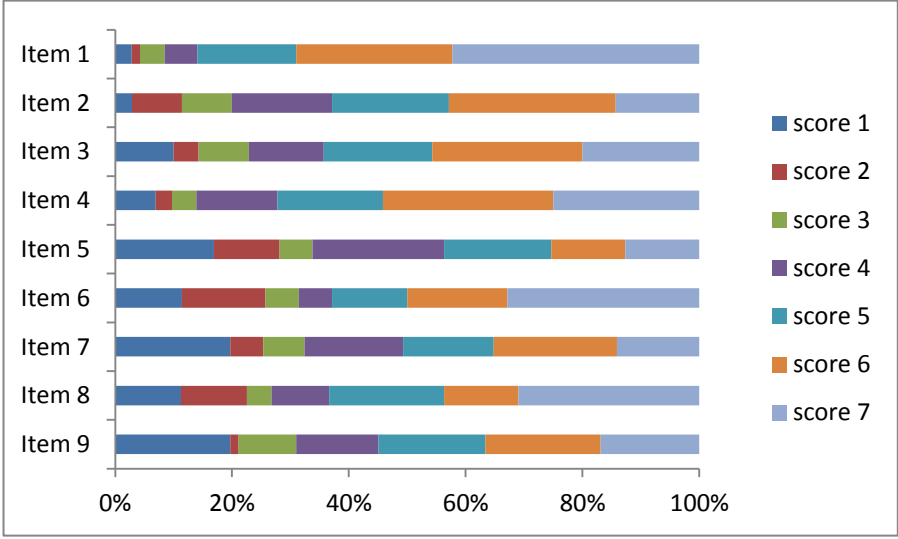


Figure 9 a. Distribution of scores on the 9 FSS items for 71 patients with congenital myopathies. The range of scores (1-7) was used in all items.

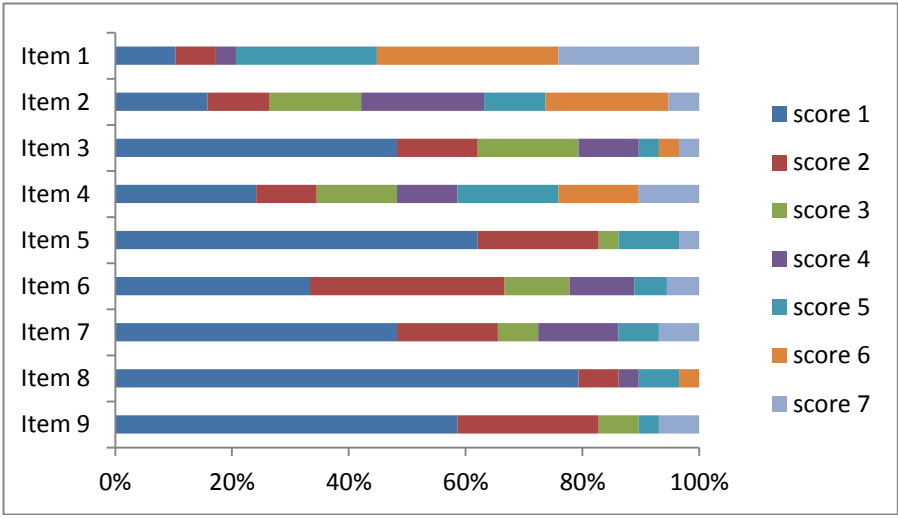


Figure 9 b. Distribution of scores on the 9 FSS items for 29 patients with SMA II. The range of scores (1-7) was used in items 2, 3 and 4. Minimum and maximum scores were used in all items except item 8. In items 5, 8 and 9 minimum score was used by $\geq 50\%$ of the patients.

The focus-group interviews indicated that the FSS in general were relevant and meaningful for patients with CM, but also that the FSS did not cover all aspects of fatigue in this group of patients. Several participants in the focus-groups commented that physical fatigue as measured by the FSS

was part of their fatigue experience, but mental fatigue was also an issue, which was not captured in the scale.

The content of item FSS 2 (“exercise brings on my fatigue”) was questioned by the CM focus-group that felt the statement was ambiguous, since exercise could bring both fatigue and energy. Item FSS 2 was apparently not applicable to patients with SMA, since 50% of the answers to this item were missing in the SMA group.

Construct validity

Correlation between FSS and VAS scores was 0.71 (Spearman’s rho) in the SMA II group and 0.59 in the CM group.

The analyses of uni-dimensionality (principal component analyses), which assess whether the same construct/component of fatigue underlies all nine FSS items, showed that the first component explained 63% of the variance in the CM group and 58% of the variance in the SMA II group. A general criterion is that the first component should explain at least 60% of the variance if the instrument is considered to be uni-dimensional. The test for consistency among FSS items (Cronbach’s alpha coefficient) was high in the CM group (0.92) as well as the SMA II group (0.90), but item-rest correlations for item FSS 1 and FSS 2 were not satisfactory. This indicated that more than one construct of fatigue was present, a finding that was also observed in the CM focus groups-interviews where the content of items FSS 1 and FSS 2 was much commented. In addition, half of the patients with SMA II had not responded to item FSS 2.

The construct and the uni-dimensionality of the FSS were improved when items FSS 1 and 2 were omitted from the scale, since the first component could now explain 72% of the variance in the CM-group and 64% of the variance in the SMA II group. Consistency, as measured by Chronbach alpha coefficients, was also slightly improved in both disease groups.

MAIN RESULTS

Study I

- Age and physical performance are related in patients with SMA II. Younger patients ≤ 20 years perform better than patients ≥ 21 years.
- The range of physical capacity in patients with SMA II is better shown in a reduced manual muscle test of the upper limbs rather than a total muscle test of the whole body.
- Hypermobility in fingers and wrist ulnar deviation is common. Limited range of motion in shoulders, elbows and wrist extension is general in patients ≥ 21 years, who also have more asymmetry in contractures compared to younger patients.
- Very limited mouth opening is a general finding in adult persons with SMA II.
- FVC% is higher in supine position compared to sitting position.

Study II

- The manual muscle test is superior to hand-held dynamometry (Citec TM) to measure physical capacity in very weak patients with SMA II, and the test can discriminate among patients with a wide range of upper limb function.
- A downscaled version of the EK2 scale “the EK2 upper limb module” and the MFM scale D3 are both equally fit to discriminate between all levels of upper limb function in patients with SMA II.
- MFM D3 is less sensitive when used entirely as an upper limb scale (without the item measuring dorsal flexion of foot).

Study III

- Muscle strength of the upper limb deteriorates over time in patients with SMA II and III; the manual muscle test recorded as MRC% can be used as method to show the decline.
- Loss of physical function over time can be demonstrated by the EK scale.
- The decline is slow and patients must be monitored over years if loss of muscle strength and physical function should be demonstrated.

Study IV

- Fatigue is a characteristic symptom in patients with CM, but not in patients with SMA II.
- Fatigue has a high impact on participation level and daily life in patients with CM.
- The Fatigue Severity Scale can be used as an instrument to capture the impact of fatigue in both disease groups.
- The scale can be used in its present form, but the scale properties and the comprehension of the scale will improve if the two first items are omitted.

DISCUSSION

In this study we have evaluated the applicability of clinical methods to assess function at impairment and activity levels in patients with SMA II, and at participation level also in patients with CM.

Patients with SMA II represent a group of non-ambulant patients with very limited muscle strength and as shown in this study, the course of the disease is slowly progressive. This puts special demands on the clinical methods used to evaluate these patients. The methods must be able to reflect all patients' level of function – from the stronger to the weaker patient - without any floor or ceiling effect; otherwise it will not be possible to register gain or loss of function as result of the disease course or of intervention. As shown, not all clinical methods evaluated in this study were equally valid to measure impairment and activity in patients with SMA II.

As described, a holistic picture of the patients and their performances can only be achieved if assessments are performed at all ICF domains of function. An international consensus on which methods should be used and which scales are appropriate has yet to be reached. The lack of consensus means that different assessment methods are used, which means that results from various studies cannot be compared. However, there is a growing international collaboration to overcome these problems and by means of networks e.g. Treat-NMD [www.treat.nmd.eu] existing scales are now being evaluated for use in clinical trials.

In the following text these issues will be discussed as will the individual methods used in this study.

Generic versus disease specific scales

In neuromuscular diseases, as in other types of diseases, generic scales are the choice at impairment level with the purpose to diagnose and to assess the patient's current status. To obtain more detailed information on the patient's function at activity and participation levels, a mixture of disease/condition specific and generic scales are used. The advantages of generic scales are that results can be directly compared with healthy populations, other disorders or between two different types of neuromuscular disorders [Streiner 2008, Hjollund 2007]. The limitation of generic scales are that the scales – especially those at activity and participation levels - may be less useful to detect minor differences and changes, and therefore become less sensitive and informative in assessing the specific disease. The disease or condition specific scales have been developed on that background, motivated from an intention to describe the characteristics and the capabilities in a particular disease

population. The disease/condition specific scale targets the condition; this is, however, not synonymous with the scale being sufficiently sensitive to include the total population. An example is that the HFMS for SMA is not suitable for the weakest patients. Although they have been developed for a specific population, some scales are used as generic scales in other populations, such as the FSS. Since the scale properties cannot automatically be transferred to another population, the use of a scale in a new population requires testing for validity and reliability in the new population.

Scale level

Measurements in clinical research is performed at different scale levels dependent on the type of variables used such as nominal, ordinal or interval variables. A great deal of information is obtained by measurements at ordinal level, such as the functional scales used in the studies of this thesis. The ordinal level of measurement means that the intervals between scale scores might differ and when scores are summed, the result will be less precise than scores obtained at an interval level of measurement. The “correct” levels of measurements have been an ongoing discussion for decades [Wright 1989, Grimby 2012] indicating that valid platforms for evaluation should be based on interval levels of measurements. Transformation of functional scales from ordinal to interval scales can be reached by Rasch Analysis, where a uni-dimensional hierarchy of scale items is built and the scale’s construct and dimensionality is assessed [Wright 1977].

Clinical methods that measure a concept at ordinal level often provide more relevant information on performance and/or perception than interval scales. There is (often) no hierarchy in the items; consequently, two individuals can have the same total score, but very different scores on the individual items. The degree of symptoms/the amount of the concept measured is based on the total score. It is important that the scale measures a single construct or that an eventual multi-construct is known. A certain amount of subjects are needed if an ordinal scale should be transformed to interval level by Rasch analysis, a criteria that is not easily fulfilled in the limited number of neuromuscular patients; furthermore some items will most likely be omitted during the Rasch procedure; items that may contain important clinical information. Consequently, the scale may be more accurate with equal intervals between scale steps, but this could be at the expense of content validity needed in a clinical setting.

Assessments at impairment level

The methods used to evaluate impairment were all on interval scales (Newton, degree, milliliters) except for the MMT in which scores were recorded as ordinal scores.

Manual muscle test

The use of MMT has been much debated. The method is a “common language” among physiotherapists and neurologists in clinical practice, but is more complicated in clinical trials when several evaluators are involved. The scales’ ordinal origin means that the distances between the scores (“0”- “5”) cannot be defined and the somewhat vague definition of the scores influences the reliability. Transforming the ordinal sum score to percentage could be dubious, since the ordinal scale is now interpreted as an interval scale. Despite this, we used the MMT in our studies as the prime method and as mentioned in the method section, we believe we fulfilled the criteria that make the method more reliable (standardized protocols, limited number of very experienced evaluators, testing weak muscles). MMT has qualities that are not met by other methods, and we showed that MMT was the only method at impairment level that could assess the capacity of all patients – also the very weak patients, a finding that is shared by others [Paternostro-Sluga 2008, Mahoney 2009]. By targeting the MMT to upper extremities a subset score of the upper limb could discriminate better among patients with various level of upper limb function than a total muscle test – both among strongest and weakest patients (fig 1) and a subset score of muscle strength in the upper limbs was adequately sensitive to register loss of muscle strength over time. A reduced muscle test of the upper limbs is less time consuming, and less troublesome for the patients with SMA II, who can remain in their wheelchair during the assessment. Furthermore, a subset of the MMT score could be more relevant, since functions are best preserved in the upper limbs and finally, reliability could be improved, since fewer muscles are tested and fewer test positions are needed. MMT subset scores has been studied and validated in other diseases [Jepsen 2004, Rider 2010], and could be further validated in SMA II. In the absence of a golden standard at impairment level, and in attempt to make the MMT more robust there is a requirement to validate the test by psychometric analysis [Cuthbert 2007]. This has been done by Vanhouette et al [2011] in a mixed group of nine different disorders (including neuromuscular disorders) based on physicians scores from eight different studies. The study concluded that the Rasch model expectations could be fulfilled by transforming the present six MRC categories (from 0 – 5) to a modified MRC sum score based on four categories: 0, paralysis; 1 severe weakness (> 50% loss of strength); 2, slight weakness (< 50% loss of strength) and 3, normal strength. This study illustrates the dilemma between the need for strong

models of measurement and the necessity of preserving the clinical estimation. Applying this shorter version of the scale to a population of patients with SMA II might improve reliability of the MMT, but the responsiveness and sensitivity of the scale would decrease, since the vast majority of our patients would be placed in the category 1.

Hand held dynamometry

Muscle strength measured by HHD is a precise and valid recording presumed an accurate initial position of the patient and the dynamometer [Merlini 2002, Delitto 1990, Jones and Stratton 2000] and performed by experienced evaluators [van der Ploug 1989 Bohannon 1988] or by a limited number of raters [Goonetilleke 1994]. These criteria were all fulfilled in our studies; however, we found that the HHD Citec™ is not applicable for use in a total population of patients with spinal muscular type II, since the dynamometer is not adequately sensitive to measure the very weak patients, in whom function can be registered by other clinical methods. Other dynamometers may have a lower threshold, but this might not improve reliability since it allows for unwanted movements to be recorded. The limitation in measuring weak patients has been found by others [Escolar 2001, Mahoney 2009] and may have the consequence that these patients cannot be enrolled in clinical trials [Miller 2001].

Range of motion

Range of motion is an essential part of physical function. Contractures may limit function, but can also improve function; as example contractures in elbow flexors may preserve the ability to lift hand to mouth since the muscles act over a shorter lever arm, but at the same time contractures in elbow flexors limit the range of reach. Similar to others, we found that contractures in the upper limbs tended to worsen with age [Wang 2004, Fajak 2010]. Furthermore, we found that contractures were related to Brooke level of upper limb function; however, a difference in sum of contractures was only found among patients at Brooke levels 3 and 4. Since joint mobility and muscle strength are both components in physical function and changes of these elements influence each other, we cannot clarify whether the loss of function from Brooke levels 3 to 4 was caused by the contractures or the contractures was a consequence of the loss of function. The fact that

contractures were more prominent in the stronger arm underlines the functional collaboration and the necessity of evaluating ROM both in clinical practice and in clinical trials.

Respiratory capacity

Respiratory function measured as FVC% is used as one out of several indicators to determine the need for supportive ventilation in spinal muscular atrophy [Mellins 1974, Manzur 2003, Wang 2007]. The method is recommended as outcome measure, but whether FVC% changes over time has been unclear. Some longitudinal studies have shown a decline over years [Carter 1995, Steffensen 2002, Kaufmann 2012] in contrast to others [Iannaccone et al 2000, Kaufmann et al. 2011] where no significant decline was shown. The various results could be related to the follow-up period, since the decline was found over a period of 5- 7 years in contrast to a period of 12 and 18 months, but could also be related to other factors. We didn't find a significant difference in FVC% between younger and older patients (≤ 20 years >), and FVC% did not change significantly over a period of 13-16 years, neither in patients with SMA II (n=10) nor in patients with SMA III (n=5). Nevertheless, two patients with SMA II ventilated via tracheostomy at the last assessment had both experienced a respiratory collapse caused by a lung infection. The finding of a relatively stable course over years is supported by earlier findings of the diaphragm being one of the best preserved muscles in SMA [Kuzuhara 1981]. Respiratory function might thus be more influenced by external factors such as lung infections rather than the natural course of disease. If so, the method might not be an appropriate outcome measure in clinical trials.

FVC% is based on the patient's respiratory performance and the recorded height. Despite the fact that the same method for measuring height was used throughout the years, measuring arm span with a flexible measuring tape could cause measurement errors, which could have influenced the recorded FVC%. However, we do not think this would have fundamental influence on the results.

Assessments at activity level

The methods used to evaluate these two ICF levels were all ordinal scales. The scales used to evaluate activity were disease-specific; the scales used to evaluate influence of fatigue on participation were not developed for neuromuscular disorders and thus used as a more generic scale.

Brooke upper limb scale

We consider the Brooke scale to be a classification rather than a scale. As a measurement scale it is too crude and has a ceiling effect among weak patients who have some residual hand function, but cannot hold a pencil or pick up coins. As classification it is easily administered and gives a quick impression on the patient's upper limb capabilities. The individual items need to be more precisely defined to ensure consistency among evaluators. In its present form, it is unclear whether the patient may use his armrest to lift the hand to the mouth in items 3 and 4. We defined these tasks without arm rest, and found excellent agreement among evaluators. By using the Brooke scale to classify our SMA II patients, we got an immediate impression of the considerable number of weak patients that did not have any function left in shoulders and elbows, but nevertheless could drive their wheelchair and operate their computer by hand.

Egen Klassifikation scale

The EK scale was developed for patients with DMD and SMA and is based on functional abilities of daily living. The scale is widely used and has been translated into several languages. It was later extended (EK2) to capture more detailed information on feeding and distal hand function [Steffensen 2008]. In this study, we used the original scale in studies I and III, and the extended version in study II. In all three studies, the sum score was used for comparison among groups, and was found to be sufficiently sensitive to discriminate within patients (over time) and between patients (younger/older and most levels of upper limb function). Neither ceiling nor floor effects were found in either of the versions. To focus on arm function we used the sum score of five EK2 upper limb items, and found that this subscale could discriminate among all patients, also among the weakest patients. This finding confirms our assumption that upper limb function must be brought into focus if variances and changes in physical function should be found in patients with SMA II.

Hammersmith functional motor scale

The HFMS is designed for children with SMA. The scale was developed from a scale designed for use in DMD, which may have influenced the choice of items that all assess gross motor function. By focusing on motor functions as sitting, rolling and changing position the weakest patients who

have lost their independent sitting balance cannot score on the scale. This results in a considerable floor effect among weak patients, and we found that almost half of our patients could not perform any task on the scale. This was not only so in adult persons, but also in weak children. The scale may be useful for some of the more physical capable patients, but it is a problem that only two items assess upper limb function, which is done at a functional level that corresponds to Brooke levels 1 and 2 (able to lift arms and bend shoulders). Consequently, the HFMS is not appropriate as an outcome measure in a total population of patients with SMA II.

Motor Function Measure

The MFM was developed for neuromuscular disorders and is as such a disease-specific scale. However, targeting all neuromuscular disorders makes the scale more generic. The various disorders are very different in origin and in phenotypes, meaning that the diseases may share similar problems at activity level, but very different problems at impairment level. As an example, two persons may have problems in walking – one because of a progressive dystrophy resulting in loss of proximal muscle strength, the other because of neuropathy resulting in loss of distal muscle strength. The scale assesses motor function in three dimensions. None of the dimensions are targeting upper limb function, but motor performance from both upper and lower limbs are included in the proximal and distal dimension. The distal dimension contains six items that assess distal function in upper limb, and only one item that assess distal function in lower limbs. We chose to omit the lower limb item to test the dimensions applicability to assess upper limb function in SMA II, and found that the subscale could then not discriminate among the strongest patients. This is likely due to the fact that none of the six items measure antigravity function of the elbow or shoulder. With such items included, the scale could properly be more sensitive and of more interest as an outcome measure for upper limb function in SMA II. Since the MFM measures the ability to perform activities that are not directly related to the patient's daily function the scale is of more limited use when the patient's daily function is assessed. However, the various tasks can be “translated” to daily activities; as an example do the ability to move one finger across squares, indicate that the patient has capacity to use electronic equipment e.g. to drive his wheelchair or use a computer.

Assessments at participation level

Fatigue Severity Scale

The FSS was developed as a disease specific scale for use in MS and SLE [Krupp 1989], but its present use in numerous disorders means that it is often considered as a generic scale. The scale has been used in other neuromuscular disorders [Drory 2001, Gagnon 2008], but has to our knowledge never been validated for use in these disorders. We showed that the scale can be used to reflect the impact of fatigue in SMA II and CM. Our analysis on the scale's uni-dimensionality resulted in findings similar to findings in other populations [Mills 2009, Ledal 2010], namely that the scale properties regarding uni-dimensionality would improve if the first two items were omitted; a finding that was supported by our qualitative data collected from two focus-group interviews. The interviews were used as means to evaluate the scale's content validity. This approach to assess content validity is recommended to assure that the concept of interest captures the concept from the patients' perspective [Rothmann 2009], and that the scale contains data that are relevant to the patient and can identify the patients' experiences [Patrick 2011]. We believe that the use of mixed methods to evaluate the properties of FSS provided results that could not have been obtained by an exclusively quantitative approach. The combined analysis of interviews and statistical calculations were an important process to understand the consistency between evidence and interpretations of the scores. The use of focus-group gains ground in medical research, also to identify the patients' priorities within research [Nierse 2013].

Visual analogue scale

VAS was used as a supplement to the FSS in the original study by Krupp [1989]. It is a generic instrument widely used to rate the intensity of a phenomenon e.g. fatigue. Whether VAS data are at interval or ordinal level is still much debated [Kersten 2012, Price 2012]; in this study we treated the scale as ordinal. The VAS scale is often used to assess FSS validity by demonstrating a degree of concurrent validity among the two scales. This assumption can be questioned since the two scales assess very different issues. The FSS assesses the impact of fatigue on daily functioning and the VAS aims at reporting the intensity of fatigue. This is not the same; a feeling of fatigue can be very intense for a period, but does not necessarily last for a long time, which could be the reason for the moderate correlations between FSS and VAS in the CM group. This finding has been observed by others [Flachenecker 2002]. The very strong correlations between FSS and VAS in the SMA II

group, do not contradict this; since fatigue was not a problem among patients with SMA II, a low score on both scales would be expected.

CONCLUSION

In this study we have made a comprehensive description of the physical characteristics in spinal muscular atrophy type II. By including the total population of Danish patients with SMA II we have shown that adults have lesser muscle strength, minor physical capabilities and more contractures than children and youngsters, and by repetitive assessments over several years we have clarified that this type of neuromuscular disease follows a slow progressive course resulting in loss of muscle strength and physical capabilities over time.

We have assessed the measurements most often used to evaluate impairment and activity in patients with SMA II, and we have emphasized that since muscle strength is best preserved in the upper limbs, it is important to relate outcomes of muscle strength and motor function to this area. At impairment level, this can be done by a MMT subscore of the upper limbs, whereas the hand held dynamometer (Citec TM) used in this study, doesn't have the ability to measure the weakest patients. At activity level, both the EK2 subscore of five items and the MFM D3 scale can differentiate among patients at all stages of SMA II. Since the EK2 scale measures the capability to perform daily activities and the MFM measures motor function, it can be recommended that the scales are used complementary. The reliability of MMT subscore and EK2 subscore needs to be further studied.

Fatigue is a factor that may influence participation. In this study, fatigue was shown to be characteristic in patients with CM, but not in patients with SMA II. The FSS, which is one of the most commonly used scales to assess the impact of fatigue was appropriate to identify and evaluate the presence of fatigue in both CM and SMA II. However, similar to other studies we also found that the scale properties would improve if the two first items were omitted. Larger groups of patients are warranted to clarify other issues of fatigue, e.g. if fatigue differs according to the individual CM types, if fatigue in SMA II is dependent on the degree of personal assistance, and what can be done to minimize fatigue in the patients.

REFERENCES

Aitkens S, Lord J, Bernauer E, Fowler WM Jr, Lieberman JS, Berck P. Relationship of manual muscle testing to objective strength measurements. *Muscle Nerve* 1989; 12(3):173-77.

American Academy of Orthopaedic Surgeons. *Joint Motion. Methods of Measuring and Recording*. Vancouver, 1965.

Arkblad E, Tulinius M, Kroksmark AK, Henriksson M, Darin N. A population-based study of genotypic and phenotypic variability in children with spinal muscular atrophy. *Acta Paediatr* 2009; 98(5):865-72.

Bäckman E, Odenrick P, Henriksson KG, Ledin T. Isometric muscle force and anthropometric values in normal children aged between 3.5 and 15 years. *Scand J Rehabil Med* 1989; 21(2):105-14.

Bäckman E, Johansson V, Hager B, Sjoblom P, Henriksson KG. Isometric muscle strength and muscular endurance in normal persons aged between 17-70 years. *Scand J Rehabil Med* 1995; 27:109-117.

Beenakker EA, van der Hoeven JH, Fock JM, Maurits NM. Reference values of maximum isometric muscle force obtained in 270 children aged 4-16 years by hand-held dynamometry. *Neuromuscul Disord* 2001; 11(5):441-46.

Berard C, Payan C, Hodgkinson I, Fermanian J. A motor function measure for neuromuscular diseases. Construction and validation study. *Neuromuscul Disord* 2005; 15(7):463-70.

Bohannon RW. Make tests and break tests of elbow flexor muscle strength. *Phys Ther* 1988; 68(2):193-94.

Boone DC, Azen SP, Lin CM, Spence C, Baron C, Lee L. Reliability of goniometric measurements. *Phys Ther* 1978; 58(11):1355-60.

Boström K, Ahlstrom G. Living with a chronic deteriorating disease: the trajectory with muscular dystrophy over ten years. *Disabil Rehabil* 2004; 26(23):1388-98.

Brooke MH, Griggs RC, Mendell JR, Fenichel GM, Shumate JB, Pellegrino RJ. Clinical trial in Duchenne dystrophy. I. The design of the protocol. *Muscle Nerve* 1981; 4(3):186-97.

Carter GT, Abresch RT, Fowler WM Jr, Johnson ER, Kilmer DD, McDonald CM. Profiles of neuromuscular diseases. Spinal muscular atrophy. *Am J Phys Med Rehabil* 1995; 74(5) Suppl: 150-159.

Cuthbert SC, Goodheart GJ Jr. On the reliability and validity of manual muscle testing: a literature review. *Chiropr Osteopat*. 2007 Mar 6; 15:4.

Delitto, A. Isokinetic dynamometry. *Muscle Nerve* 1990; 13 Suppl:S53-S57.

de Vries JM, Hagemans ML, Bussmann JB, van der Ploeg AT, van Doorn PA. Fatigue in neuromuscular disorders: focus on Guillain-Barre syndrome and Pompe disease. *Cell Mol Life Sci* 2010; 67(5):701-13.

Deymeer F, Serdaroglu P, Parman Y, Poda M. Natural history of SMA IIIb: muscle strength decreases in a predictable sequence and magnitude. *Neurology* 2008; 71(9):644-49.

Drory VE, Goltsman E, Reznik JG, Mosek A, Korczyn AD. The value of muscle exercise in patients with amyotrophic lateral sclerosis. *J Neurol Sci* 2001; 191 (1-2):133-37.

Dubowitz V. Chaos in the classification of SMA: a possible resolution. *Neuromuscul Disord* 1995; 5(1):3-5.

Dubowitz, V. Disorders in the Lower Motor Neurone: the Spinal Muscular Atrophies. *Muscle Disorders in Childhood*. London: W.B. Saunders Company, 1995. 325-69.

The 184th ENMC workshop "Pain and fatigue in NMD: prevalence and management" Naarden, the Netherlands, 20-22th May 2011.

European medicines Agency. Reflection Paper on the Regulatory Guidance for the Use of health Related Quality of Life (HRQL) Measures in the Evaluation of Medicinal Products. London: European Medicines Agency 2005.

Escolar DM, Henricson EK, Mayhew J, Florence J, Leshner R, Patel KM et al. Clinical evaluator reliability for quantitative and manual muscle testing measures of strength in children. *Muscle Nerve* 2001; 24(6):787-93.

Feasson L, Camdessanche JP, El Mandhi L, Calmels P, Millet GY. Fatigue and neuromuscular diseases. *Ann Readapt Med Phys* 2006; 49(6):289-84.

Feldkotter M, Schwarzer V, Wirth R, Wienker TF, Wirth B. Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *Am J Hum Genet* 2002; 70(2):358-68.

Flachenecker P, Kumpfel T, Kallmann B, Gottschalk M, Grauer O, Rieckmann P et al. Fatigue in multiple sclerosis: a comparison of different rating scales and correlation to clinical parameters. *Mult Scler* 2002; 8(6):523-26.

Florence JM, Pandya S, King WM, Robison JD, Signore LC, Wentzell M et al. Clinical trials in Duchenne dystrophy. Standardization and reliability of evaluation procedures. *Phys Ther* 1984; 64(1):41-45.

Florence JM, Pandya S, King WM, Robison JD, Signore LC, Wentzell M. Intrarater reliability of manual muscle test (Medical Research Council scale) grades in Duchenne's muscular dystrophy. *Phys Ther* 1992; 72(2):115-22.

Folio MR and Fewell RR. *Peabody Developmental Motor Scales. Second Edition*. Ed. TX. PRO-ED: Austin. 2000.

Fowler WM Jr. Rehabilitation management of muscular dystrophy and related disorders: II. Comprehensive care. *Arch Phys Med Rehabil* 1982; 63(7):322-28.

Fujak A, Kopschina C, Gras F, Forst R, Forst J. Contractures of the upper extremities in spinal muscular atrophy type II. Descriptive clinical study with retrospective data collection. *Ortop Traumatol Rehabil* 2010; 12(5):410-19.

Gajdosik RL, Bohannon RW. Clinical measurement of range of motion. Review of goniometry emphasizing reliability and validity. *Phys Ther* 1987; 67(12):1867-72.

- Gagnon C, Mathieu J, Jean S, Laberge L, Perron M, Veillette S et al. Predictors of disrupted social participation in myotonic dystrophy type 1. *Arch Phys Med Rehabil* 2008; 89(7):1246-55.
- Goonetilleke A, Modarres-Sadeghi H, Guilloff RJ. Accuracy, reproducibility, and variability of hand-held dynamometry in motor neuron disease. *J Neurol Neurosurg Psychiatry* 1994; 57(3):326-32.
- Grimby G, Tennant A, Tesio L. The use of raw scores from ordinal scales: time to end malpractice? *J Rehabil Med* 2012; 44(2):97-98.
- Hagemans ML, van Schie SP, Janssens AC, van Doorn PA, Reuser AJ, van der Ploeg AT. Fatigue: an important feature of late-onset Pompe disease. *J Neurol* 2007; 254(7):941-45.
- Heatwole C, Bode R, Johnson N, Quinn C, Martens W, McDermott M.P et al. Patient-reported impact of symptoms in myotonic dystrophy type 1 (PRISM-1). *Neurology* 2012; 79(4):348-57.
- Henderson, S. E., & Sugden, D. A. *Movement Assessment Battery for Children*. London: Psychological Corporation 1992.
- Hepper NG, Black LF, Fowler WS. Relationships of Lung Volume to Height and Arm Span in normal subjects and in patients with spinal deformity. *Am Rev Respir Dis* 1965; 91:356-62.
- Hjollund NH, Andersen JH, Bech P. Assessment of fatigue in chronic disease: a bibliographic study of fatigue measurement scales. *Health Qual Life Outcomes* 2007; 5:12.
- Hsieh H, Shannon SE. Three approaches to qualitative content analysis. *Qualitative Health Research* 2005; 15(9):1277-1288.
- Iannaccone ST, Russman BS, Browne RH, Buncher CR, White M, Samaha FJ. Prospective analysis of strength in spinal muscular atrophy. DCN/Spinal Muscular Atrophy Group. *J Child Neurol* 2000; 15(2):97-101.
- Jepsen J, Laursen L, Larsen A, Hagert CG. Manual strength testing in 14 upper limb muscles: a study of inter-rater reliability. *Acta Orthop Scand* 2004; 75(4):442-48.
- Jones, Stratton G. Muscle function assessment in children. *Acta Paediatr* 2000; 89(7):753-61.

Jungbluth H, Zhou H, Hartley L, Halliger-Keller B, Messina S, Longman C et al. Minicore myopathy with ophthalmoplegia caused by mutations in the ryanodine receptor type 1 gene. *Neurology* 2005; 65 (12):1930-35.

Kalkman JS, Schillings ML, van der Werf SP, Padberg GW, Zwarts MJ, van Engelen BG et al. Experienced fatigue in facioscapulohumeral dystrophy, myotonic dystrophy, and HMSN-I. *J.Neurol.Neurosurg.Psychiatry* 2005; 76(10):1406-09.

Kaufmann P, McDermott MP, Darras BT, Finkel R, Kang P, Oskoui M, et al. Observational study of spinal muscular atrophy type 2 and 3: functional outcomes over 1 year. *Arch Neurol* 2011; 68(6):779-86.

Kaufmann P, McDermott MP, Darras BT, Finkel RS, Sproule DM, Kang PB et al. Prospective cohort study of spinal muscular atrophy types 2 and 3. *Neurology* 2012; 79(18):1889-97.

Kersten P, Kucukdeveci AA, Tennant A. The use of the Visual Analogue Scale (VAS) in rehabilitation outcomes. *J Rehabil Med* 2012; 44(7):609-10.

Kleinman L, Zodet M W, Hakim Z, Aledort J, Barker C, Chan K et al. Psychometric evaluation of the Fatigue Severity Scale for use in chronic hepatitis C. *Quality of Life Research* 2000; 9(5):499-508.

Kleyweg RP, van der Meche FG, Schmitz PI. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barre syndrome. *Muscle Nerve* 1991; 14(11):1103-09.

Kroksmark AK, Beckung E, Tulinius M. Muscle strength and motor function in children and adolescents with spinal muscular atrophy II and III. *Eur J Paediatr Neurol* 2001; 5(5):191-98.

Krosschell KJ, Maczulski JA, Crawford TO, Scott C, Swoboda KJ. A modified Hammersmith functional motor scale for use in multi-center research on spinal muscular atrophy. *Neuromuscul Disord* 2006; 16(7):417-26.

Krueger RA and Casey MA. *Focus Groups: A Practical Guide for Applied Research*. 4th ed. Sage, 2008.

Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol* 1989; 46(10):1121-23.

Kuzuhara S, Chou SM. Preservation of the phrenic motoneurons in Werdnig-Hoffmann disease. *Ann Neuro* 1981; 9(5):506-10.

Laberge L, Gagnon C, Jean S, Mathieu J. Fatigue and daytime sleepiness rating scales in myotonic dystrophy: a study of reliability. *J Neurol Neurosurg Psychiatry* 2005; 76(10):1403-05.

Lee KA, Hicks G, Nino-Murcia G. Validity and reliability of a scale to assess fatigue. *Psychiatry Res* 1991; 36(3):291-98.

Lerdal, A.; Johansson, S.; Kottorp, A.; von, Koch L. Psychometric properties of the Fatigue Severity Scale: Rasch analyses of responses in a Norwegian and a Swedish MS cohort. *Mult Scler* 2010; 16(6):733-41.

Lou JS, Weiss MD, Carter GT. Assessment and management of fatigue in neuromuscular disease. *Am J Hosp Palliat Care* 2010; 27(2):145-57.

Lovett RV, Martin EG. Certain aspects of infantile paralysis with a description of a method of manual muscle testing. *J Am Med Assoc* 1916; 66: 729-733.

Lyager S., Steffensen B, Juhl B. Indicators of need for mechanical ventilation in Duchenne muscular dystrophy and spinal muscular atrophy. *Chest* 1995; 108(3):779-85.

Mahony K, Hunt A, Daley D, Sims S, Adams R. Inter-tester reliability and precision of manual muscle testing and hand-held dynamometry in lower limb muscles of children with spina bifida. *Phys Occup Ther Pediatr* 2009; 29(1):44-59.

Mafi P, Mafi R, Hindocha S, Griffin M, Khan W. A Systematic Review of Dynamometry and its Role in Hand Trauma Assessment. *Open Orthop J* 2012; 6: 95–102.

Main M, Kairon H, Mercuri E, Muntoni F. The Hammersmith functional motor scale for children with spinal muscular atrophy: a scale to test ability and monitor progress in children with limited ambulation. *Eur J Paediatr Neurol* 2003; 7(4):155-59.

Manzur AY, Muntoni F, Simonds A. Muscular dystrophy campaign sponsored workshop: recommendation for respiratory care of children with spinal muscular atrophy type II and III. 13th February 2002, London, UK. *Neuromuscul Disord* 2003; 13(2):184-89.

Mazzone E, Bianco F, Martinelli D, Glanzman AM, Messina S, De Sanctis R et al. Assessing upper limb function in nonambulant SMA patients: development of a new module. *Neuromuscul Disord* 2011; 21(6):406-12.

McCormack P, Woods S, Aartsma-Rus A, Hagger L, Herczegfalvi A, Heslop E et al. Guidance in social and ethical issues related to clinical, diagnostic care and novel therapies for hereditary neuromuscular rare diseases: "translating" the translational. *PLoS Curr* 2013 jan 10;5.

Medical Research Council: Aids to investigation of peripheral nerve injuries. London: Her Majesty's Stationary Office; 1943

Mellins RB, Hays AP, Gold AP, Berdon WE, Bowdler JD. Respiratory distress as the initial manifestation of Werdnig-Hoffmann disease. *Pediatrics* 1974; 53(1):33-40.

Mercuri E, Messina S, Battini R, Berardinelli A, Boffi P, Bono R et al. Reliability of the Hammersmith functional motor scale for spinal muscular atrophy in a multicentric study. *Neuromuscul Disord* 2006; 16(2):93-98.

Merlini L, Mazzone ES, Solari A, Morandi L. Reliability of hand-held dynamometry in spinal muscular atrophy. *Muscle Nerve* 2002; 26(1):64-70.

Miller F, Koreska J. Height measurement of patients with neuromuscular disease and contractures. *Dev Med Child Neurol* 1992; 34(1):55-60.

Miller RG, Moore DH, Dronsky V, Bradley W, Barohn R, Bryan W. A placebo-controlled trial of gabapentin in spinal muscular atrophy. *J Neurol Sci* 2001; 191(1-2):127-31.

Mills R, Young C, Nicholas R, Pallant J, Tennant A. Rasch analysis of the Fatigue Severity Scale in multiple sclerosis. *Mult Scler* 2009; 15(1): 81-87.

Munsat TL, Davies KE. International SMA consortium meeting. (26-28 June 1992, Bonn, Germany). *Neuromuscul Disord* 1992; 2(5-6):423-28.

Nelson L, Owens H, Hynan LS, Iannaccone ST. The gross motor function measure is a valid and sensitive outcome measure for spinal muscular atrophy. *Neuromuscul Disord* 2006; 16(6):374-80.

Nierse CJ, Abma TA, Horemans AM, van Engelen BG. Research priorities of patients with neuromuscular disease. *Disabil Rehabil* 2013; 35(5):405-12.

North K. What's new in congenital myopathies?. *Neuromuscul Disord* 2008; 18(6):433-42.

Paternostro-Sluga T, Grim-Stieger M, Posch M, Schuhfried O, Vacariu G, Mittermaier C et al. Reliability and validity of the Medical Research Council (MRC) scale and a modified scale for testing muscle strength in patients with radial palsy. *J Rehabil Med* 2008; 40(8):665-71.

Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E et al. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1--eliciting concepts for a new PRO instrument. *Value Health* 2011; 14(8):967-77.

Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E et al. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2--assessing respondent understanding. *Value Health* 2011; 14(8):978-88.

Pettersson S, Moller S, Svenungsson E, Gunnarsson I, Welin Henriksson E. Women's experience of SLE-related fatigue: a focus group interview study. *Rheumatology (Oxford)* 2010; 49(10):1935-42.

Price DD, Staud R, Robinson ME. How should we use the visual analogue scale (VAS) in rehabilitation outcomes? II: Visual analogue scales as ratio scales: an alternative to the view of Kersten et al. *J Rehabil Med* 2012; 44(9):800-01.

Rider LG, Koziol D, Giannini EH, Jain MS, Smith MR, Whitney-Mahoney K et al. Validation of manual muscle testing and a subset of eight muscles for adult and juvenile idiopathic inflammatory myopathies. *Arthritis Care Res (Hoboken)*. 2010; 62(4):465-72.

Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report. *Value Health* 2009; 12(8):1075-83.

Russman BS, Buncher CR, White M, Samaha FJ, Iannaccone ST. Function changes in spinal muscular atrophy II and III. The DCN/SMA Group. *Neurology* 1996; 47(4):973-76.

Russman BS, Melchreit R, Drennan JC. Spinal muscular atrophy: the natural course of disease. *Muscle Nerve* 1983; 6(3):179-81.

Ryan MM, Schnell C, Strickland CD, Shield LK, Morgan G, Iannaccone ST et al. Nemaline myopathy: a clinical study of 143 cases. *Ann Neurol* 2001; 50(3):312-20.

Samaha FJ, Buncher CR, Russman BS, White ML, Iannaccone ST, Barker L et al. Pulmonary function in spinal muscular atrophy. *J Child Neurol* 1994; 9(3):326-29.

Saraniti AJ, Gleim GW, Melvin M, Nicholas JA. The relationship between subjective and objective measurements of strength *. *J Orthop Sports Phys Ther* 1980; 2(1):15-19.

Schillings ML, Kalkman JS, Janssen HM, van Engelen BG, Bleijenberg G, Zwarts MJ. Experienced and physiological fatigue in neuromuscular disorders. *Clin Neurophysiol* 2007; 118(2):292-300.

Scott OM, Hyde SA, Goddard C, Dubowitz V. Quantitation of muscle function in children: a prospective study in Duchenne muscular dystrophy. *Muscle Nerve* 1982; 5(4):291-301.

Steffensen B, Hyde SA, Lyager S, Mattsson E. Validity of the EK-scale: a functional assessment of non-ambulatory individuals with Duchenne muscular dystrophy or spinal muscular atrophy. *Physiother Res Int* 2001; 6(3):119-34.

Steffensen B, Mayhew A, Aloysius A. Egen Klassifikation (EK) revisited in spinal muscular atrophy. *Neuromuscul Disord* 2008; 8:740-41.

Steffensen BF, Lyager S, Werge B, Rahbek J, Mattsson E. Physical capacity in non-ambulatory people with Duchenne muscular dystrophy or spinal muscular atrophy: a longitudinal study. *Dev Med Child Neurol* 2002; 44(9):623-32.

Stratford PW, Balsor BE. A comparison of make and break tests using a hand-held dynamometer and the Kin-Com. *J Orthop Sports Phys Ther* 1994; 19(1):28-32.

Streiner DL, Norman GR. *Health Measurement scales. A practical guide to their development and use.* 4th ed. Oxford: Oxford University Press. 2008

Swoboda KJ, Prior TW, Scott CB, McNaught TP, Wride MC, Reyna SP et al. Natural history of denervation in SMA: relation to age, SMN2 copy number, and function. *Ann Neurol* 2005; 57(5):704-12.

Thieme A, Mitulla B, Schulze F, Spiegler AW. Epidemiological data on Werdnig-Hoffmann disease in Germany (West-Thuringen). *Hum Genet* 1993; 91(3):295-97.

Tseng BY, Gajewski BJ, Kluding PM. Reliability, responsiveness, and validity of the visual analog fatigue scale to measure exertion fatigue in people with chronic stroke: a preliminary study. *Stroke Res Treat* 2010; 2010. <http://www.ncbi.nlm.nih.gov/pubmed/20700421>

United Nations. *Standard Rules on the Equalization of Opportunities for Persons with Disability*, adopted by the United Nations General Assembly at its 48th session on the 20 december 1993 (resolution 48/96). New York, NY, United Nations Department of Public Information, 1994.

US Food and Drug Administration Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Rockville, MD: Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, 2009.

Valko PO, Bassetti CL, Bloch KE, Held U, Baumann CR. Validation of the Fatigue Severity Scale in a Swiss cohort. *Sleep* 2008; 31(11):1601-07.

van der Ploeg RJ, Oosterhuis HJ. The "make/break test" as a diagnostic tool in functional weakness. *J Neurol Neurosurg Psychiatry* 1991; 54(3):248-51.

Vanhoutte EK, Faber CG, van Nes SI, Jacobs BC, van Doorn PA, van Koningsveld R. Modifying the Medical Research Council grading system through Rasch analyses. *Brain* 2012; 135(Pt 5):1639-49.

Vignos PJ Jr, Spencer GE Jr, Archibald KC. Management of progressive muscular dystrophy in childhood. *JAMA* 1963; 184:89-96.

Wade DT. Outcome measures for clinical rehabilitation trials: impairment, function, quality of life, or value? *Am J Phys Med Rehabil* 2003; 82(10 Suppl):S26-S31.

Wadsworth CT, Krishnan R, Sear M, Harrold J, Nielsen DH. Intrarater reliability of manual muscle testing and hand-held dynamometric muscle testing. *Phys Ther* 1987; 67(9):1342-47.

Wang CH, Finkel RS, Bertini ES, Schroth M, Simonds A, Wong B et al. Consensus statement for standard of care in spinal muscular atrophy. *J Child Neurol* 2007; 22(8):1027-49.

Wang CH, Bonnemann CG, Rutkowski A, Sejersen T, Bellini J, Battista V et al. Consensus statement on standard of care for congenital muscular dystrophies. *J Child Neurol* 2010; 25(12):1559-81.

Wang CH, Dowling JJ, North K, Schroth MK, Sejersen T, Shapiro F et al. Consensus statement on standard of care for congenital myopathies. *J Child Neurol* 2012; 27(3):363-82.

Wang HY, Ju YH, Chen SM, Lo SK, Jong YJ. Joint range of motion limitations in children and young adults with spinal muscular atrophy. *Arch Phys Med Rehabil* 2004; 85(10):1689-93.

Whitbeck C. Four Basic Concepts of Medical Science. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1978:210-222.

WHO definition of Health. Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.

WHO International Classification of Functioning, Disability and Health: World Health Organisation, Geneva 2001.

Wirth B, Brichta L, Schrank B, Lochmuller H, Blick S, Baasner A et al. Mildly affected patients with spinal muscular atrophy are partially protected by an increased SMN2 copy number. *Hum Genet* 2006; 119(4):422-28.

Wright BD. Solving measurement problems with the Rasch model. *Journal of Educational Measurement* 1977; 14(2):97-116.

Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil* 1989; 70(12):857-60.

Wokke JHJ. Fatigue is a part of the burden of neuromuscular diseases. *J Neurol* 2007; 254(7):948-949.

Zerres K, Rudnik-Schoneborn S. Natural history in proximal spinal muscular atrophy. Clinical analysis of 445 patients and suggestions for a modification of existing classifications. *Arch Neurol* 1995; 52(5):518-23.

APPENDIX – THE FUNCTIONAL SCALES, PAPERS

Brooke upper limb scale (Brooke et al. 1981)

Item	
1	Starting with the arms at the sides the patient can abduct arms in full circle until they touch above head
2	Can raise arms above head only by flexing the elbow
3	Cannot raise arms above head but can raise 8 oz glass of water to mouth
4	Can raise hands to mouth but cannot raise 8 oz glass of water to mouth
5	Cannot raise hands to mouth but can use hands to pick up pennies from table
6	Cannot raise hands to mouth and has no useful function of hands

Motor Function Measure D3 (Berard et al. 2005)

Distal dimension. Each item is scored from 0 – 3 with higher score representing higher function. Item number refers to the number on the MFM.

Item	
4	From plantar flexion, dorsiflexes the foot to at least 90° in relation to leg. (Supine, leg supported by examiner)
17	10 coins on table – successively picks up and holds 10 coins in hand during 20-second period
18	One finger placed in center of fixed CD – goes round the edge of CD with one finger without contact of the hand on the table
19	Pencil on table. Pick up the pencil and draw loops inside frame
20	Holding sheet of paper: tears the sheet of paper
21	Tennis ball on table : pick up the ball and turn the hand over completely holding the ball
22	One finger placed in diagram (nine squares): raise the finger and place it successively on the squares

Egen Klassifikation (EK) scale (Steffensen et al. 2001)

Egen Klassifikation Scale Version 2 (EK2)		
Steffensen 2008		
1	Ability to use wheelchair. How do you get around indoors and outdoors?	N/A
	Able to use a manual wheelchair on flat ground, 10m < 1 minute	0
	Able to use a manual wheelchair on flat ground, 10m > 1 minute	1
	Unable to use manual wheelchair, requires power wheelchair	2
	Uses power wheelchair, but occasionally has difficulty steering	3
2	Ability to transfer from wheelchair. How do you transfer from your wheelchair to a bed?	
	Able to transfer from wheelchair without help	0
	Able to transfer independently from wheelchair, with use of aid	1
	Needs assistance to transfer with or without additional aids (hoist, easy glide)	2
	Needs to be lifted with support of head when transferring from wheelchair	3
3	Ability to stand. Do you sometimes stand? How do you do this?	
	Able to stand with knees supported, as when using braces	0
	Able to stand with knees and hips supported, as when using standing aids	1
	Able to stand with full body support	2
	Unable to be stood	3
4	Ability to balance in the wheelchair. Can you bend forwards and to the sides and return to the upright position?	
	Able to push himself upright from complete forward flexion by pushing up with hands	0
	Able to move the upper part of the body > 30 in all directions from the upright position, but cannot push himself upright as above	1
	Able to move the upper part of the body < 30 from one side to the other	2
	Unable to change position of the upper part of the body, cannot sit without total support of the trunk and head	3
5	Ability to move the arms. Can you move your fingers, hands and arms against gravity?	
	Able to raise the arms above the head with or without compensatory movements	0
	Unable to lift the arms above the head, but able to raise the forearms against gravity, ie. hand to mouth with / without elbow support	1
	Unable to lift the forearms against gravity, but able to use the hands against gravity when the forearm is supported	2
	Unable to move the hands against gravity but able to use the fingers	3
6	Ability to use the hands and arms for eating. Can you describe how you eat?	
	Able to eat and drink without elbow support	0
	Eats or drinks with support at elbow	1
	Eats and drinks with elbow support; with reinforcement of the opposite hand +or – aids	2
	Has to be fed	3
7	Ability to turn in bed. How do you turn in bed during the night?	
	Able to turn himself in bed with bedclothes	0
	Needs some help to turn in bed or can turn in some directions	1
	Unable to turn himself in bed. Has to be turned 0 - 3 times during the night	2
	Unable to turn himself in bed. Has to be turned > 4 times during the night	3
8	Ability to cough. How do you cough when you have to?	
	Able to cough effectively	0
	Has difficulty to cough and sometimes needs manual reinforcement. Able to clear throat	1
	Always needs help with coughing. Only possible to cough in certain positions	2
	Unable to cough, Needs suction and/or hyperventilation techniques or IPPB in order to keep airways clear	3

9	Ability to speak. Can you speak so that what you say can be understood if you sit at the back of a large room?	N/A
	Powerful speech. Able to sing and speak loudly	0
	Speaks normally, but cannot raise his voice	1
	Speaks with quiet voice and needs a breath after 3 to 5 words	2
	Speech is difficult to understand except to close relatives	3
10	Physical well-being. (This relates to respiratory insufficiency only – see manual)	
	No complaints, feels good	0
	Easily tires. Has difficulty resting in a chair or in bed	1
	Has loss of weight, loss of appetite, Scared of falling asleep at night, sleeps badly	2
	Experience additional symptoms: change of mood, stomach ache, palpitations, perspiring,	3
11	Daytime fatigue. Do you have to organise your day or take a rest to avoid getting too tired?	
	Doesn't get tired during day	0
	Need to limit activity to avoid getting too tired	1
	Need to limit my activity and have a rest period to avoid getting too tired	2
	Get tired during day even if I rest and limit activity	3
12	Head Control. How much head support do you need in your wheelchair?	
	Does not need head support	0
	Needs head support when going up and down slope (15° standard ramp)	1
	Needs head support when driving wheelchair	2
	When sitting still in a wheelchair needs head support	3
13	Ability to control Joystick. What kind of joystick do you use to control your chair?	
	Uses a standard joystick without special adaptation	0
	Uses an adapted joystick or has adjusted wheelchair in order to use joystick	1
	Uses other techniques for steering than joystick such as blowing sucking systems or scanned driving	2
	Unable to operate wheelchair. Needs another person to operate it	3
14	Food Textures. Do you have to modify your food in any way in order to eat it?	
	Eats all textures of food	0
	Eats cut up food or avoids hard/chewy foods	1
	Eats minced/ pureed food	2
	Main intake consists of being tube fed	3
15	Eating a meal. (with or without assistance) How long does it take to complete a meal?	
	Able to consume a whole meal in the same time as others sharing the meal	0
	Able to consume a whole meal in the same time as others only with encouragement or needs some additional time	1
	Able to consume a whole meal but requires substantially more time compared to others (15 m or more extra)	2
	Unable to consume a whole meal	3
16	Swallowing. Do you ever have problems with swallowing?	
	Never has problems when swallowing and never chokes on food/drink,	0
	May experience occasional (less than once a month) problems swallowing certain types of food or occasionally chokes	1
	Has regular trouble swallowing food/drink or chokes on food/drink (more than once a month)	2
	Has trouble swallowing saliva or secretions	3
17	Hand function. Which of these activities can you do?	
	Can unscrew the lid of a water or fizzy drink bottle and break the seal	0
	Can write two lines or use computer keyboard	1
	Can write signature or send text or use remote control	2
	Cannot use hands	3
	Total	

The Hammersmith functional motor scale for SMA (HMFS) (Main et al. 2003)

Score 2	Score 1	Score 0
Frog/chair sitting no hand support	One hand support	Two hand support
Long sitting, no hands	One hand support	Two hand support
1/2 roll from supine, both ways	One way (R/L?)	Unable
Touches one hand to head (R/L?) (in sitting)	Flexes head to hand	Unable
Touches two hands to head (in sitting)	Flexes head to hands	Unable
Rolls prone to supine over R	Pushes on hand	Unable
Rolls prone to supine over L	Pushes on hand	Unable
Rolls supine to prone over R	Pulls on hand	Unable
Rolls supine to prone over L	Pulls on hand	Unable
Gets to lying from sitting (safely)		Unable
Achieves prop on forearms-head up	Holds position when placed	Unable
Lifts head from prone (arms down by sides)		Unable
Achieves four point kneeling-head up	Holds position when placed	Unable
Achieves prop on extended arms-head up	Holds position when placed	Unable
Gets to sitting from lying through side lying	Through prone	Unable
Crawls	Crawls 2 m	Unable
Lifts head from supine	Through side flexion	Unable
Stands holding on with one hand	Stands with minimal trunk support	Hip/knee support needed
Stands independently count ≥ 3	Stands independently count < 4	Stands momentarily
Takes four steps unaided	Takes 2–4 steps unaided	Unable


Total

Fatigue Severity Scale (FSS) (Krupp et al. 1989)

Your Name _____ Date: _____

This questionnaire contains nine statements that rate the severity of your fatigue symptoms. Read each statement and circle a number from 1 to 7, based on how accurately it reflects your condition during the past week and the extent to which you agree or disagree that the statement applies to you.

***A low value (e.g. 1) indicates strong disagreement with the statement, whereas a high value (e.g. 7) indicates strong agreement.

During the past week, I have found that:	Disagree						Agree
1. My motivation is lower when I am fatigued	1	2	3	4	5	6	7
2. Exercise brings on my fatigue.	1	2	3	4	5	6	7
3. I am easily fatigued.	1	2	3	4	5	6	7
4. Fatigue interferes with my physical functioning.	1	2	3	4	5	6	7
5. Fatigue causes frequent problems for me.	1	2	3	4	5	6	7
6. My fatigue prevents sustained physical functioning.	1	2	3	4	5	6	7
7. My fatigue interferes with carrying out certain duties and responsibilities.	1	2	3	4	5	6	7
8. Fatigue is among my three most disabling symptoms.	1	2	3	4	5	6	7
9. Fatigue interferes with my work, family or social life.	1	2	3	4	5	6	7

Total Score: _____