Evaluation of scales and measurement instruments in immune-mediated polyneuropathies

I.S.J. Merkies

ISBN 90-9014393-9

Evaluation of scales and measurement instruments in immune-mediated polyneuropathies

Evaluatie van meetmethoden in immuungemedieerde polyneuropathiëen

Evaluashón di métodonan pa midi den polyneuropatía immuno-mediá

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. dr. ir. J.H. van Bemmel
en volgens besluit van het College voor Promoties.
de openbare verdediging zal plaatsvinden op
woensdag 31 januari 2001 om 11.45 uur

door

Ingemar Sergio José Merkies

geboren te Willemstad, Curaçao, Nederlandse Antillen

Promotiecommissie

Promotor: Prof. dr. F.G.A. van der Meché

Overige leden: Prof. dr. R.A.C. Hughes
Prof. dr. P.J. Koudstaal
Prof. dr. J. Passchier

Copromotoren: Dr. ir. P.I.M. Schmitz
Dr. P.A. van Doorn

The studies described in this thesis were performed at the Department of Neurology, University Hospital Rotterdam/Erasmus University, Rotterdam, The Netherlands.

This research project was financially supported by the European Community: Biomed project number: BMH4-CT96 0324, Brussels.

The realisation of this thesis was supported by grants from Serono Benelux, Janssen-Cilag, GlaxoWellcome, Smith-Kline Beecham Farma, Biogen International, NASKHO, PricewaterhouseCoopers, Brakkeput Mei-Mei, and the Stichting Neuro-immunologisch onderzoek AZR.

Printed by PrintPartners 'Ipskamp', Enschede, The Netherlands.

Aan mijn dochter, echtgenote, familie & vrienden
Aan mijn ouders, dat ze in vrede moge rusten
God zegen ons & ons eiland

CONTENTS page

Chapter

Chapter 1

General introduction

1.1 Immune-mediated polyneuropathies – Clinical aspects

1.2 Evaluating outcome measures
   A. International Classification of Impairments, Disabilities, and Handicaps (ICIDH)
   B. Quality of life concept
   C. Clinical appropriateness of outcome measures: Simplicity, Validity, Reliability, Responsiveness, Communicability

1.3 Evaluating outcome measures in immune-mediated polyneuropathy clinical studies

1.4 Purpose and design of Inflammatory Neuropathy Cause And Treatment (*INCAT*) group study.

1.1 Immune-mediated polyneuropathies – Clinical aspects

Introduction

Immune-mediated neuropathies mainly include Guillain-Barré syndrome (GBS) (most common form: acute inflammatory demyelinating polyneuropathy (AIDP)), chronic inflammatory demyelinating polyneuropathy (CIDP), neuropathy associated with monoclonal gammopathy of undetermined significance (MGUSP), and multifocal motor neuropathy (MMN) (1-11). Electrophysiological examination in these patients generally reveals features of a demyelinating polyneuropathy. These neuropathies have become increasingly important in clinical neurology because they are readily diagnosable and potentially treatable. Evidence has emerged from many papers in the last decade to indicate that these illnesses represent part of a continuum separated by their neuromuscular dysfunction pattern, time course and response to various treatments (Figure 1; 1-11). However, the distinction between these illnesses is in some aspect somewhat artificial. In the following, selected aspects of these diseases are discussed with particular emphasis on clinical presentation and outcome measures applied in clinical studies published from January 1988 to January 1999 that included patients with one of these disorders.

Guillain-Barré syndrome (GBS)

The clinical diagnosis of Guillain-Barré syndrome (GBS) is generally not difficult to establish. GBS diagnosis is based on criteria originally developed to aid epidemiological field studies (12,13). In 1993, the World Health Organisation has postulated new diagnostic criteria for this illness (14). The criteria for GBS have been broadened and are currently entirely based on clinical features (14). Symmetrical weakness and decreased or disappearance of the myotatic reflexes are the major criteria. Surprisingly, sensory disturbances are considered supportive features, despite the fact that most patients with GBS experience a sensory-motor pattern (1,15). Acute sensory polyneuritis as a variant form of GBS has also been described (16).
Cerebrospinal fluid analysis is of limited value. The protein concentration rises usually from the second week onward. A high cell count is usually not found, but one should consider HIV-related GBS or Borreliosis in the infrequent cases with pleiocytosis. Electromyography helps to characterise the pattern of GBS as either more demyelinating or axonal (17).
From a clinical perspective, it is generally known that a large interindividual variability exists in GBS. The limitations of the presented diagnostic criteria are most obviously reflected when weakness does not start in the extremities but in the cranial nerves such as in Miller Fisher syndrome, which is characterised by oculomotor paresis, ataxia, and areflexia, or in the lower bulbar variant, characterised by difficulties in speech and swallowing. Confluent patterns have also been demonstrated.

GBS can be distinguished from CIDP by a different time course. By definition, the nadir of GBS is reached within 4 weeks, but in the majority of patients within 2 weeks (1,12,13). For CIDP, in contrast, duration of progressive weakness of at least 2 months has been suggested as a criterion (18). More recently, a group of patients have been described with an intermediate "subacute" time course (SIADP; with a nadir between 4 to 8 weeks of onset) (19). In fact, large series showed that there is a continuum between GBS and CIDP with a sharp peak of patients who have their nadir in the first few weeks and a long tail of patients who have a much longer and usually less fulminate course (20,21).

Chronic inflammatory demyelinating polyneuropathy (CIDP)

Chronic inflammatory demyelinating polyneuropathy (CIDP) is progressive over a period of more than 2 months (18). Thereafter weakness may progress or may be stable during months or years, or the patients may improve spontaneously followed by a relapsing-remitting course (22). Patients with CIDP have symmetrical distal more than proximal weakness that generally predominates over sensory deficit. Sensory-motor form is most commonly seen, although pure motor or pure sensory patterns have been reported (23,24). Areflexia is common. Cerebrospinal fluid generally shows an increased protein without a cellular reaction. Electrophysiological criteria for CIDP have been proposed and include evidence for demyelination with features such as conduction blocks, dispersion or the compound muscle action potential, increased distal latencies, or slowed conduction velocities (25). Prior to making the diagnosis of CIDP, it is essential to rule out other causes of chronic polyneuropathies. Diseases such as hereditary neuropathies, vasculitis, cryoglobulinemia,

multiple myeloma, and many others have to be excluded. Dyck and associates stated that the clinical diagnosis of CIDP could only be made in the absence of a systemic disease (26). However, it seems that CIDP can also occur in the setting of some concurrent diseases (4,27). For example, patients with an otherwise typical clinical picture compatible with CIDP have been described having an IgG monoclonal gammopathy of undetermined significance (8,27).

Polyneuropathy associated with monoclonal gammopathy of undetermined significance (MGUSP)

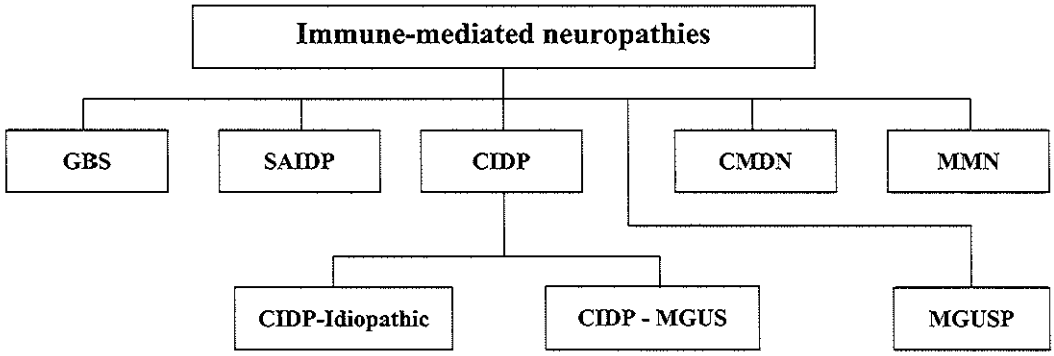
At present it is still a matter of debate whether the diagnosis of CIDP can be made when a monoclonal protein of undetermined significance (MGUS) is present. Patients with a neuropathy completely compatible with the clinical diagnosis of CIDP may have a MGUS. Several studies have compared patients with an idiopathic CIDP with those associated with an MGUSP (28-30). Up to 25-30% of patients with a clinical diagnosis of CIDP also had a MGUS (28-30). In general, patients with a CIDP-MGUS are at onset of symptoms older demonstrating a more smouldering clinical course and on average experience more frequent sensory loss with less severe weakness, despite similar motor conduction findings when compared with CIDP patients without MGUS. However, application of these clinical differences should be done with some caution, because these differences have not been demonstrated in a more recent study (23).

Multifocal Motor Neuropathy (MMN)

Multifocal motor neuropathy (MMN) is an infrequent occurring chronic immune-mediated demyelinating neuropathy (31-33). Patients with MMN mostly have a stepwise progression of asymmetrical muscle weakness and amyotrophy localised in the anatomical distribution areas of peripheral nerves. Sensory symptoms are generally absent. Patients with MMN may experience cramps, fasciculation or myokymia. The electrophysiological hallmark of MMN is persistent motor conduction block with reduction of the motor nerve conduction velocity only over the affected areas. Clinically, MMN is also described as an asymmetrical pure motor variant of chronic inflammatory demyelinating polyneuropathy (CIDP) with multifocal motor conduction blocks. Especially, during the evolution of MMN the multifocal character may gradually evolve in a more or less symmetrical pattern, clinically resembling the motor form of CIDP. Neuropathological studies have also linked MMN to CIDP (34,35). A multifocal sensory-motor demyelinating polyneuropathy have also been reported, hence fulfilling the intermediate clinical pattern between CIDP and MMN (Figure 1; 36,37).

Figure 1

Clinical spectrum of immune-mediated neuropathies



Legend to Figure 1
Guillain-Barré syndrome (GBS) (also generally known as acute inflammatory demyelinating polyneuropathy (AIDP)) is characterised by a monophasic course of weakness that reaches a nadir within four weeks, the majority within two weeks followed by a plateau-phase with gradual recovery hereafter. GBS is primarily distinguished from chronic inflammatory demyelinating polyneuropathy (CIDP) by a different time course. For CIDP the duration of progressive weakness is at least 2 months. Weakness prevails over sensory deficit. More recently, a group of patients have been described with an intermediate "subacute" time course (SIADP; nadir between 4 to 8 weeks of onset) (19). Patients with a polyneuropathy associated with a monoclonal gammopathy of undetermined significance (MGUSP) have generally more sensory than motor deficit. Immunoglobulin G MGUSP may have great resemblance with CIDP. Multifocal motor neuropathy (MMN) is characterised by asymmetric motor deficit, affecting the arms more than the legs with no significant sensory abnormalities. MMN is considered the asymmetrical motor variant of CIDP. An intermediate form, the chronic multifocal sensory-motor demyelinating polyneuropathy (CMDN) have been recently described, thus completing the clinical spectrum of the most common forms of inflammatory polyneuropathies.

1.2 Evaluating outcome measures

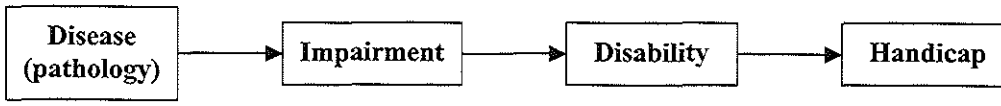
A. International Classification of Impairments, Disabilities, and Handicaps (ICIDH)

The first step in the clinical evaluation of patients is to determine at which level outcome is going to be assessed and which scale or set of scales is going to be used for this purpose. This is particularly complex in neurology, because the nervous system has so many functions, whereas most other organs have a much more limited range of functions. As an example, Figure 1 depicts the spectrum of inflammatory disorders of the peripheral nerves that can range from pure symmetrical or asymmetrical weakness to pure sensory multi-modality disturbances with respiratory problems in GBS cases. Therefore, the first question to answer in evaluating a patient is what information should be measured and how this is going to be obtained. Gathering information on particular clinical features is generally done by firstly reviewing the literature on existing outcome measures and secondly classifying selected measure(s) according to the postulated international classification of impairments,

disabilities, and handicaps (ICIDH) by the World Health Organisation (WHO) in 1980 (Figure 2; 38). The ICIDH model is widely used and was developed as an attempt to standardise classification and terminologies relating to the consequences of health conditions. According to this model, any disease (pathology) can be evaluated at the following defined levels:

Figure 2

The International Classification of Impairments, Disabilities, and Handicaps (ICIDH)



1. *Disease* (*pathology*) refers to the damage or abnormal processes occurring within an organ or organ system inside the body

2. *Impairment* is defined as the direct physiological consequences of the underlying pathology. In other words, impairment is defined as any loss or abnormality of psychological, physiological, or anatomical structure or function. Impairment represents exteriorisation of a pathological state, and as such it represents a disturbance at the organ level.

3. *Disability* is defined as any restriction or lack (resulting from impairment) of ability to perform an activity in the manner or within the range considered normal for a human being. Disability represents objectification of an impairment, and as such represents disturbances at the level of the person.

4. *Handicap* represents socialisation of an impairment or disability, and as such it reflects the consequences for the individual - cultural, social, economic, and environmental - that stem from the presence of impairment and disability. In other words: This is the disadvantage for a given individual, resulting from an impairment or a disability, that limits or prevents the fulfilment of a social role that is normal (depending on age, sex, and social and cultural factors) for that individual.

Although this classification seems to be straightforward, these levels form a continuum with no clear-cut boundaries. Despite this, this concept is extremely useful for purposes of evaluating and classifying outcome measures, and therefore it is important to understand the technical definition of each of these levels. Also, it is important to note that not all impairment features may lead to disability. Examples of the latter are: 1) the detection of conduction blocks in a nerve without any functional deficit in daily activities or a completely recovered patient with CIDP with normal daily functionality that still has areflexia at neurological examination. Also, impairment can sometimes lead directly to handicap. This is the case in a patient with residual ophthalmoplegia due to Miller-Fisher syndrome who is not able to work as a bus-driver anymore. Another example is a positive test for HIV that

can cause loss of insurance, social isolation, etcetera, without any intervening impairment and disability.

B. Quality of life concept

Assessing outcome using the ICIDH levels is primarily being performed by clinicians. However, the obtained results are strongly dependent on several patients' factors, both intra-individual and environmental. Examples of such factors are individual ability of patients to adapt to illness (coping), factors intended to increase functioning such as rehabilitation, physical therapy training, patient's perseverance, and amount of social support. Therefore, outcome can also be studied from the "patient's own perspective", a concept captured in "quality of life" outcome measures. Quality of life is also defined as the patient's reaction to the discrepancy between actual and expected achievements arising as a consequence of illness (39,40). At least four dimensions should be included in a quality of life assessment. These dimensions are physical, functional, psychological, and social health. The physical health dimension refers primarily to disease-related and treatment-related symptoms. Functional health comprises self-care, mobility, and physical activity level, as well as the capacity to carry out various roles in relation to family and work. Cognitive functioning, emotional status, and general perceptions of health, well-being, life satisfaction, and happiness are the central components of the psychological life domain. Social functioning includes the assessment of qualitative and quantitative aspects of social contacts and interactions (39,40).

C. Clinical appropriateness of outcome measures: Simplicity, Validity, Reliability, Responsiveness, Communicability

To identify scales of potential interest for measuring a specific feature, clinicians have to find their way through a daunting array of available scales. Subsequently, because there is an increasing emphasis on accuracy, an important step in choosing a scale comprises a critical review of its clinical appropriateness to the patient group at study and evaluation of its clinimetric requirements (40-44). A useful scale should fulfil the following requirements (40-44):

1. It should be simple, none-time consuming with little or no special training
2. It should be valid
3. It should be reliable
4. It should be responsive to changes over time in the underlying condition, yet relatively insensitive to symptoms or signs fluctuations
5. It should provide results that easily can be interpreted by others

1. *It should be simple, none-time consuming with little special training.* Many outcome measures are impractical because they require too much time to administer or score. A measure should wherever possible be simple, particularly if more than one person is

going to use the measure. Simplicity will improve the patient (and user) compliance, and will increase reliability.

2. *It should be valid.* A valid scale is one that measures what it purports to measure and therefore provides the information required. In other words: It should accurately describe the underlying phenomenon or disease. Various types of validity are described:

   Face validity refers to the apparent sensibility of the measure and its components. It simply indicates whether, on the face of it, the scale appears to be assessing the desired qualities. This validity form represents the subjective judgement based on a review of the measure itself by one or more experts, and rarely are any empirical approaches used. The functional grading scale (f-score), used in many GBS studies, has obvious face validity for the assessment of mobility (45).

   Content validity is closely related to the face validity concept. It consists of a judgement by experts evaluating whether an outcome measure captures all the relevant or important contents or domains of an illness. The f-score would be a valid measure for lower limb function in neuropathies, because it is strongly based upon mobility (45).

   These two forms of validity are also entitled as "the validity forms by assumption", meaning that a measure assesses outcome in a certain way because an expert says it does (46).

   Construct validity is demonstrated by examining the relations between a newly created test and other tests to show that the new test measures the same 'construct'. In practice, evidence for construct validity is gathered by undertaking a series of studies to determine:

   Convergent validity – the extent to which a measure correlates with other measures of related entities.
   Discriminant validity – the extent to which a measure does not correlate with measures of different entities.
   Divergent validity – the extent to which a measure correlates with measures of opposite entities. Others have not described this construct validity form. However, we introduce this validity form as part of the validation of rating scales. It is a matter of debate whether the proposed 'divergent validity' should be considered as a 'discriminant validity' form. An example: The correlation between fatigue and vitality scales.

   Criterion-related validity is demonstrated by examining the accuracy of a test compared with a particular standard, the criterion ('gold standard'). There are two types of criterion-related validity: concurrent and predictive. The distinction between the two refers to whether the measure is compared with a gold standard measured at the same time (concurrent) or in the future (predictive).

Concurrent validity – the extent to which a new measure correlates with another widely accepted validated measure or the opinion of experts. This is generally applied in cases of no "real" gold standard.
Predictive validity – if we thought that the f-score values at 4 weeks of follow-up in patients with GBS could predict degree of disability at 6 months, data collected at these two instances in time must be correlated (45).

3. *It should be reliable.* A reliable measure is one that produces results that are accurate, consistent, stable over time, and reproducible. A patient whose condition has not changed should always receive the same score apart from random variation. There are three different types of reliability:

   Internal consistency (Interitem consistency) is the extent to which items comprising a scale measure the same concept – that is, a measure of the homogeneity of the scale. There are a number of ways to calculate these correlations, of which the Cronbach's alpha is the most widely used (47).

   Observer reliability is the agreement between observers or within an individual observer. There are two types:

   Interobserver reliability is the agreement between observations made by two or more raters on the same patient or group of patients.
   Intraobserver reliability is the agreement between observations made by the same rater on two different occasions on the same patient or group of patients.

   Test-retest reliability is the agreement between observations made by the same patient on two different occasions.

The concepts of *validity* and *reliability* can be explained using the example of "shooting at a target". Someone learning archery must first learn to hit the centre of the target, and then to do this consistently. The *validity* of a measure would be represented by the aim of shooting – how close, on average, the shots come to the centre of the target (good validity meaning a bias of approximately zero distance). The *reliability* of a measure would be represented by how close successive shots fall to each other, wherever they land on the target (good reliability meaning a small variance).

4. *It should be responsive to changes over time in the underlying condition, yet relatively insensitive to minor symptoms or signs fluctuations.* Whereas validity and reliability form the clinimetric core stones of a rating scale, the ability of a measure to detect clinically meaningful changes over time is crucial. For clinicians and researchers, such a measure should discriminate between irrelevant changes (normal fluctuations in the activity of an illness; "noise") and clinically meaningful changes on which a treatment policy can be based ("signal"), an ability addressed as "responsiveness". A statistic and heuristic approach in examining responsiveness of a measure has been proposed (43). Statistical responsiveness captures the ability of an instrument to measure any change, irrespective

of its relevance. Heuristic techniques are based upon comparing changes as assessed by a scale with an external indicator, for example the grades of judgement by the patients of their clinical condition (e.g. grade 1: improved; grade 2: stable; grade 3: deteriorated).

5. *It should provide results that easily can be interpreted by others.* A measure should give results that are easily understood by others.

Selection of a scale or set of scales is one of the most important steps in planning clinical follow-up of patients or research projects. Whatever scale is chosen, it is important that the above-mentioned requirements are fulfilled before its general use as an outcome measure.

1.3 Evaluating outcome measures in immune-mediated polyneuropathy clinical studies

Introduction

In the last decade, the assessment of disease activity in clinical studies including patients with GBS, CIDP or MGUSP has been commonly based on a multitude of measures, ranging from the assessment of general strength, sensory disturbances, and functional abilities. Various easily applicable scales have been devised and used for these purposes.
As postulated in section 1.2, each outcome measure can be classified according to the model devised by the World Health Organisation regarding the general consequences of any illness, the ICIDH which reflects three stages: Impairments, Disabilities, and Handicaps (see section 1.2; 38). In addition, outcome can also be assessed from the patient's own perspective using the so-called 'quality of life' surveys (39,40). Subsequently, clinimetric requirements need to be fulfilled before the use of any outcome measure.

Methods

To evaluate outcome in immuned-mediated polyneuropathies, a Medline search with reference tracing was conducted for the period January 1988 till January 1999 evaluating all clinical outcome measures (besides neurological examination) applied in clinical studies including patients with GBS, CIDP or MGUSP. The outcome measures were systematically categorised according to the ICIDH and Quality of life concepts. Reports published in English that included 10 patients or more were identified using the following keywords: GBS, CIDP, acquired/idiopathic (poly)(radiculo)neuropathy, polyneuritis, gammopathy, dysimmune, paraprotein(a)emia, and monoclonal gammopathy of undetermined significance polyneuropathy. Subsequently, we investigated whether the applied scales have been formally evaluated in terms of being valid, reliable, and responsive before their use as neurological outcome measures.

This search was performed on behalf of the Inflammatory Neuropathy Cause And Treatment (*INCAT*) group, a collaborating force of European neurologists with special interest in immune-mediated neurological disorders.

Medline search

Phase I *– recruitment of clinical immune-mediated polyneuropathy studies.*
Eighty-six clinical studies were collected that included at least 10 patients with GBS, CIDP or MGUSP. To strive for clarity, patients with subacute inflammatory demyelinating polyneuropathy or a GBS variant (such as Miller Fisher's syndrome) were categorised under the GBS heading. The outcome measures applied in these studies and their corresponding ICIDH level or quality of life status are presented Table 1.
At the impairment level, various motor scales have been devised and used to assess strength. These scales are primarily focused on the Medical Research Council grading system (48). The most widely used motor scales are the MRC sumscore as described by Kleyweg and associates and the motor subset of the Neurologic disability scale (NDS) (49-51). Despite its confusing name, the NDS represents outcome at the impairment level.
Sensory deficit has been assessed using different sensory rating scales that included various sensation modalities representing different sensory fibres (Table 1;15,50,51,59,62,67,70,76, 90,115,117,121).
At the disability level, two scales have been regularly applied: The Hughes' disability scale and corresponding modifications and the (modified) Rankin scale (Table 1; 45,49,52). Unfortunately, these two scales share the characteristics of being strongly directed towards mobility. Surprisingly, thus far no disability measure has been used that provides a "true general outcome", (defined as functional arms + legs information)" in patients with immune-mediated polyneuropathies.
No pure handicap scale has been applied in patients with GBS, CIDP or MGUSP and only two studies have assessed outcome from the patients' perspective using the Sickness Impact Profile health survey (53,59). One study used the Environmental Status scale to assess handicap in 10 patients with GBS (131). Unfortunately, this scale was not conceptualised according to the ICIDH guidelines and mixes disability with handicap measures (132).
Finally, some scales are devised by a mixture of items representing various ICIDH levels, thus hampering their applicability and communicability (Table 1;55,88,98,99,112,116,120).

Phase II *– literature evaluation regarding clinimetric properties of the most widely used outcome measures in immune-mediated clinical studies.*
As stated, the most commonly used scales in immune-mediated polyneuropathy studies are the MRC sumscore, the NDS, different sensory rating scales including various sensation modalities representing different sensory fibres, the Hughes' disability scale and related modifications and the (modified) Rankin scale (Table 1). Surprisingly, despite the wide use of these scales, studies formally evaluating their clinimetric requirements are limited and incomplete. The MRC sumscore and the motor subset of the NDS are the only motor scales that has been validated and examined in terms of their reproducibility (49-51). With the exception of the sensory subset of the NDS, none of the used sensory scales has been

submitted to a comprehensive clinimetric evaluation. The validity and reliability of the NDS sensory subset were demonstrated in diabetes patients with signs of a polyneuropathy (50,51). Good internal consistency was recently obtained for this scale in patients with hereditary and a variety of other polyneuropathies (53). The NDS sensory subset is, however, limited because sensory qualities are only assessed at the index finger and hallux.
Despite the wide use of the Hughes' disability scale and the Rankin scale in neuropathy studies and their obvious simplicity and face validity, the clinical value of these scales are limited because they are strongly directed towards mobility without providing information regarding the arms. The validity and reliability of the Hughes' disability scale were demonstrated in patients with GBS (49). No formal clinimetric evaluation of the Rankin scale has been performed in patients with immune-mediated polyneuropathies. Its reliability, however, was established in patients with stroke (52).
Although validity and reliability have been demonstrated for some of the scales applied in patients with GBS, CIDP or MGUSP, relative little attention have been addressed towards their ability to detect clinically meaningful changes over time (43,44).

Conclusions

Based on these observations, the following conclusions were stated regarding clinical studies including patients with immune-mediated polyneuropathies in the last twelve years:

- There is a daunting array of available outcome measures in these disorders leading to great confusion
- Some outcome measures are ill-constructed and therefore not generally applicable
- No uniformity exists regarding which scale or set of scales represents best the impairment and disability outcome levels that also covers the whole range of patients with GBS, CIDP and MGUSP
- Thus far, no "global disability" outcome measure have been applied in these disorders
- Scales that purely measure handicap have not been applied in these conditions
- Outcome from a patients' perspective using quality of life measures have been scarcely used
- Studies evaluating the clinimetric requirements of the outcome measures most applied in these disorders are limited and incomplete.

Table 1

Clinical studies reported from January 1988 to January 1999 that included at least ten patients with Guillain-Barré syndrome (GBS), chronic inflammatory demyelinating polyneuropathy (CIDP), or polyneuropathy associated with a monoclonal gammopathy of undetermined significance (MGUSP)

Clinical studies	References	Total number of enrolled patients	Impairment		Disability		Handicap	Quality of life
			Motor scales	Sensory scales	(Predominantly) Mobility scales	Other/global scales		
GBS	61,63,66,68, 69,71-73,79, 81,83,87,89,106	882	*	*	-	-	-	-
	15	100	*	A	-	-	-	-
	21,55,58,60,64, 74,75,77,80,82, 84,85	1268	*	*	(Modified) Hughes scale Functional test limbs[2]	-	-	-
	49,57,65,86,107	363	MRC sumscore[1] Mean weakness score[4]	-	(Modified) Hughes scale	-	-	-
	53	29	MRC sumscore[1] NSS/NDS	NSS/NDS	-	-	-	SIP-Ph[16]
	56,76,78	724	-	B	(Modified) Hughes scale	Disability arm grade[3]	-	-
	59	123	-	C	(Modified) Hughes scale	-	-	SIP[16]
	62	42	-	D	(Modified) Hughes scale	-	-	-
	67,70	306	MRC sumscore[1]	E	(Modified) Hughes scale	-	-	-
	88	52	-	-	Expanding grading scale[5]	-	-	-
	131	10	MRC muscle testing scale[17]	Pain-VAS Score**	Barthel Index / Rivermead-ADL[18]	Enviromental Status scale[19]	-	-
CIDP	21,105	134	*	*	Modified Hughes scale	-	-	-
	23,29,30	191	Average MRC score[15]	*	(Modified) Rankin	-	-	-
	24	10	-	NDS	-	-	-	-
	28,92,104, 113	116	*	*	-	-	-	-
	53	9	MRC sumscore[1] NSS/NDS	NSS/NDS	-	-	-	SIP-Ph[16]
	90	16	MRC scoring system[6] Grip strength	F	(Modified) Rankin	-	-	-
	91,93,94,108,11 1	288	*	*	(Modified) Rankin	-	-	-
	95	60	Average muscle score[7]	*	-	-	-	-
	96,110	80	MRC sumscore[1]	-	(Modified) Rankin	-	-	-
	98	16	*	*	Functional disability score[8]	-	-	-
	99,100	48	NDS Grip strength	NDS	Functional clinical grading scale[9]	-	-	-
	109	10	*	*		RMI[10]	-	-
	112	93	*	*	Clinical disability score[11]	-	-	-
	114	20	NDS	NDS; Case-IV	-	-	-	-

Table 1 (continues)

Clinical studies	References	Total number of enrolled patients	Impairment		Disability		Handicap	Quality of life
			Motor scales	Sensory scales	(Predominantly) Mobility scales	Other/global scales		
MGUSP	23,29,30	66	Average MRC score[15]	*	(Modified) Rankin	-	-	-
	28,122,125	132	*	*	-	-	-	-
	53	2	MRC sumscore[1] NSS/NDS	NSS/NDS	-	-	-	SIP-Ph[16]
	97,101,102, 118,119	207	NDS	NDS; Case-IV	-	-	-	-
	103,124,126-130	193	*	*	-	-	-	-
	115	36	Average MRC score[15]	G	(Modified) Rankin	-	-	-
	116,120	90	*	*	Disability severity scale[12]	-	-	-
	117,121	48	MRC scoring system[13]	H: Ataxia tapping test Vibrameter	(Modified Rankin)	-	-	-
	123	11	MRC scoring system[14]	NDS	-	-	-	-

Legend to Table
References using the same outcome measures were gathered.

Impairment level
*Routine neurological examination; - examination not performed

Motor assessment methods
[1]MRC sumscore of six muscle pairs (see reference 49); [4]Mean weakness score = mean score of 16 muscles including two proximal and two distal muscles in the arms and legs; [6]MRC sumscore of ten muscle groups; [7]Average muscle-score: of 34 muscle groups; [13]MRC sumscore of 12 muscle pairs; [14]MRC sumscore of 30 muscle pairs; [15]Average MRC score of four muscle groups (see reference 23); NSS = Neuropathy symptom score (see references 50,51); NDS = Neurologic disability scale (see references 50,51); [17]MRC muscle testing scale = sumscore of the following muscle groups: shoulder abductors, elbow flexors and extensors, wrist flexors and extensors, extensor digitorum communis, first dorsal interosseous, abductor pollicis brevis, hip flexors, knee flexors and extensors, ankle dorsiflexors and plantar flexors, extensor and flexor hallucis longus (see reference 131).

Sensory assessment methods
A. Sensory symptoms and signs were subjectively scored: 0 = absent; 1 = mild; 2 = severe (see reference 15).
B. Poorest sensory ability: 0 = normal; 1 = symptoms but no signs; 2 = anaesthesia or analgesia of fingers or feet; 3 = anaesthesia or analgesia to elbows or knees or worse (see reference 76).
C. Sensory grading system: disturbed sensation toes/fingers = 1; in the feet and hands = 2; in the legs and arms = 3 (see reference 59).
D. Sensory grading system: 0 = normal; 1 = symptoms without objective sensory loss; 2 = loss of light touch or pain sensation on fingers or toes; 3 = sensory loss to wrists or ankle; 4 = sensory loss to elbows or knees; 5 = sensory loss to shoulders or groin (see reference 62).
E. Two-point discrimination at digit II, proprioception and tactile functions in the hands and feet were assessed. A two-point discrimination value of ≥ 5 mm was considered abnormal. All qualities were scored: 0 = normal; 1 = abnormal (total score: 0 [normal sensation] to 10 [maximum sensory deficit]) (see references 67,70).
F. Lower limb sensory testing (vibration, joint position, pinprick): 3 = normal; 2 = impaired at the great toe; 1 = impaired at the ankle; 0 = impaired at or above the calf (total score: 0 - 18) (see reference 90).

G. Sensory grading system (vibration, joint position, pinprick): 3 = normal; 2 = impaired at the fingertip or big toe; 1 = impaired at the proximal interphalangeal joint or at the ankle; 0 = impaired proximal to the proximal interphalangeal joint or above the ankle (see reference 115).
H. Sensory grading system: For both touch and pinprick sense: 4 = normal; 3 = abnormal distal to wrist/ankle; 2 = abnormal distal half forearm/leg; 1 = abnormal distal to elbow/knee; 0 = abnormal distal to axilla/groin. Vibration sense was graded: 4 = tuning fork perception (128 Hz) on middle finger/hallux; 3 = ulnar styloid/medial malleolus; 2 = elbow/knee; 1 = clavicula/crista iliaca; 0 = no perception. Joint position sense of middle finger/hallux was graded: 2 = normal; 1 = diminished; 0 = absent (Total score: 0 to 56) (see references 117,121).

NSS = Neuropathy symptom score (see references 50,51).
NDS = Neurologic disability scale (see references 50,51).
Pain-VAS = Pain visual analogue scale.

Disability level

[3]0 = normal; 1 = minor symptoms or signs but able to put hand on top of head when sitting with head upright and able to oppose thumb to each fingertip; 2 = able to do either of the tasks in 1 but not both; 3 = some movements but unable to perform either of the tasks in 2; 4 = no movements; 5 = dead (see references 56,76).
[10]RMI = Rivermead mobility index (see reference 109 for description).
[18]Rivermead-ADL = Rivermead activity of daily living scale (see reference 40 for description).

Mixed level (= containing impairment and disability items)

[2]Functional tests of upper (move fingers, hold a pen between thumb and forefinger, flex forearm over arm in the pronation and supination position, lift elbow above bed plane, and maintain arms outstretched) or the lower limbs (move toes, ankle dorsiflexion and plantar flexion, knee flexion above bed plane, stand up, walk with and without assistance, and stand up from a squatting position) (see reference 55).
[5]Expanding grading scale: 0 = healthy; 1 = minimal signs or symptoms without motor involvement; 2 = involvement of cranial nerves only; 3 = minor signs or symptoms at the extremities; 4 = able to walk without support; 5 = able to walk with support; 6 = unable to walk but no complete tetraparesis and/or need for ventilation; 7 = requiring ventilation but no complete tetraparesis; 8 = complete tetraparesis without need for ventilation; 9 = complete tetraparesis and need for ventilation; 10 = dead (see reference 88).
[8]Functional disability score: 0 = normal neurological examination and normal functional state; 1 = areflexia, with or without subjective symptoms and normal functional state; 2 = neurological findings other than areflexia, without or with only mild limitation of normal function; 3 = neurological findings other than areflexia with moderate or severe limitation or function but with self-ambulation; 4 = neurological findings other than areflexia with need for a wheel-chair without self-ambulation (paraplegia) (see reference 98).
[9]Functional clinical grading scale (see reference 99).
[11]Clinical disability score: 1 = mild motor or sensory symptoms and signs; 2 = moderate motor or sensory involvement; 3 = severe involvement requiring assistance for eating, dressing, or walking (see reference 112).
[12]Disability severity scale: 0 = normal; 1 = clinical or electrophysiological signs (or both) or neuropathy without symptoms of neuropathy (subclinical neuropathy); 2 = mild motor or sensory symptoms (or both) with or without mild functional impairment; 3 = moderately disabled by motor and sensory symptoms including ataxia; 4 = requiring assistance in eating, dressing, or using a walking device; 5 = not ambulatory (see references 116,120).

Mixed level (= containing disability and handicap items)

[19]Environmental Status scale mixes disability and handicap items and is not conceptualised according to the ICIDH guidelines (132).

Quality of life level

[16]SIP = sickness impact profile (see reference 54). SIP-Ph = Physical dimension of the SIP.

Purpose and design of Inflammatory Neuropathy Cause And Treatment *(INCAT)* study

Introduction

The Inflammatory Neuropathy Cause And Treatment *(INCAT)* group is a collaboration of European neurologists that is particularly interested in the causes and evaluation of treatment-effects of inflammatory polyneuropathies. Since the *INCAT*-group has the need to evaluate the clinical effect of treatment in patients with inflammatory neuropathies, the disability level of measuring treatment-outcome is of preferential interest. At present, there is a strong need for outcome measures that fulfil all clinimetric requirements and cover the whole of inflammatory neuropathies. The Rotterdam *INCAT*-centre has been given the task to evaluate existing scales regarding their applicability in patients with an inflammatory polyneuropathies (GBS, CIDP, MGUSP) and if necessary to construct new scales to evaluate treatment outcome. This study was part of the Biomed-project, number BMH4-CT96 0324 that was supported by the European community.
An extensive review of the literature was primarily conducted (see section 1.3) and from this a set of scales was presented to the *INCAT* group (Table 2 + appendix I – IV, pages 205-221). Also, two new scales ("*INCAT*" sensory sumscore and the Rotterdam 9 Items Handicap scale) were created using a judgmental approach based on literature review, patients' suggestions, and experts' opinions to close the existing gaps in covering the whole immune-mediated clinical range (46,133). All *INCAT* members gave their opinion regarding these scales in general and on their different items. The scales were modified according to the suggestions made by the *INCAT* members and eventually face and content validity were achieved. However, further clinimetric evaluation of the presented scales was needed, because the majority of these scales have not been properly evaluated in patients with immune-mediated polyneuropathies. A study was therefore performed evaluating the clinical appropriateness of the selected measures. The study had the following objectives:

Objectives

- To evaluate the clinimetric properties (validity, reliability, responsiveness) and clinical applicability of the selected scales in patients with GBS, CIDP, and MGUSP.

- To collect normal values in healthy individuals for the following instruments: The Rydel-Seiffer graduated tuning fork for assessing vibration sense at various sites of examination and the hand-held Martin Vigorimeter for assessing grip strength (134,135).

- To evaluate the degree of fatigue in patients with immune-mediated polyneuropathies and healthy controls using the Fatigue severity scale (136).

- To investigate the clinical feasibility of the newly created scales ("*INCAT*" sensory sumscore and Rotterdam 9 items handicap scale).

Table 2

Scales and gadgets at the various levels of outcome selected for the INCAT study and reported clinimetric evaluation thus far

Outcome measures	Validity	Reliability		Responsiveness	References	Study type
		Inter-observer	Intra-observer Test-retest*			
Impairment level						
MRC sumscore	+	+	+	descriptive	49	GBS
Martin Vigorimeter	+	+	+	+	137-140	RA HC
Rydel-Seiffer graduated Tuning fork	+	+	+	descriptive	141-144	Diabetic polyneuropathy HC
"INCAT" sensory sumscore	-	-	-	-	-	
Fatigue severity scale	+	-	+*	+	136	MS Lyme SLE
Disability level						
Hughes' disability scale (f-score)	+	+	+	Descriptive	49	GBS
Arm disability scale	+	+	+	+	145	MS
Leg disability scale	+	+	+	+	145	MS
Overall Disability sumscore (Arms+Leg disability)	+	+	+	+	145	MS
Nine-hole peg test	+	+	+	-	40	Stroke
Ten-metre walking test	+	+	+	-	40	Stroke
Handicap level						
Rotterdam 9 items handicap scale	-	-	-	-	-	
'Modified Rankin'	+	+	+	-	40,52	Stroke
Quality of life survey						
SF-36	+	+	+	+	146-149	Various diseases HC

Legend to Table 2
+ = performed; - = not performed; RA = rheumatoid arthritis; MS = multiple sclerosis; HC = healthy controls; GBS = Guillain-Barré syndrome; SLE = systemic lupus erythematosus.

- To determine which scale represents best the impairment and disability levels of outcome.
- To examine the relationship between scales representing the various levels of outcome.
- To present a comprehensive evaluation of the various "responsiveness-techniques" for the selected scales.

Study setting and population

Location

- University hospital Rotterdam, department of neurology.

Patients

- Transversal group of patients (stable group): 113 patients with a stable clinical condition were recruited from the Rotterdam immune-mediated polyneuropathy databank and from the Dutch GBS study group (stable group). This group consisted of 83 patients who experienced GBS many years before the start of the study, 22 patients with CIDP, and 8 patients with a MGUSP.

- Longitudinal group of patients (longitudinal group): 20 consecutive patients with sensory-motor GBS (n = 7) or CIDP (n = 13) were recruited for responsiveness investigation of the selected scales. During recruitment period, no patients at our department were diagnosed having a MGUSP.

- A transversal group of 59 patients with various forms of minor polyneuropathy were recruited and examined in the first study that evaluated the simplicity and validity of the Rydel-Seiffer graduated tuning fork. The aetiologies of the polyneuropathies were diabetes mellitus (11 cases), hereditary motor sensory neuropathy type I (5 cases) and type II (2 cases), systemic disease related (3 cases), amyloidosis (3 cases), drug induced (2 cases), thyroid dysfunction (1 case), and vitamine B12 deficiency (1 case). Fourteen patients had a chronic idiopathic axonal polyneuropathy. In 17 cases no cause was determined.

Healthy individuals

- Regarding the investigation of the normal values of the Rydel-Seiffer graduated tuning fork: Healthy controls were recruited from hospital personnel, companions (relatives, friends) of patients visiting our outpatient clinic and from homes for the elderly. Two hundred and fourteen potential controls were interviewed and examined.

- Regarding normal values evaluation of the Vigorimeter: Five hundred and fifty-one potential healthy controls were recruited from hospital personnel, relatives and friends accompanying patients at our outpatient clinic, and healthy elderly, living in the village "Nieuw-Vennep" as part of the municipal of Haarlemmermeer, The Netherlands. There was a door-to-door mailing and an article was published in a local newspaper explaining the purpose and significance of the study. Attempts were made to obtain participants representing a wide variety of social and occupational backgrounds. The children and adolescents were recruited at one primary and one high school.

- Regarding evaluation of the degree of fatigue in healthy controls: One hundred and thirteen age and gender matched healthy controls were recruited from hospital personnel, companions (relatives, friends) of patients visiting our outpatient clinic and from homes for the elderly.

Study design and plan

Patients
Each patient was interviewed before the start of the study by the main-investigator (IM) to collect general characteristics (age, gender, diagnosis, date of diagnosis, which hospital admitted in the past if applicable, course of illness).

Healthy controls
For the normal values evaluation of the Rydel-Seiffer graduated tuning fork and Vigorimeter, healthy controls were stratified for age and gender. Fatigue was evaluated in the stable group of patients and the controls (n = 113) were age and gender matched.

Investigators
Two senior neurologists (PD, FM) and six residents in Neurology (IM, JS, FO, RM, WM, and RK) formed 28 different pairs of examiners (Table 3). "Experienced" couple (couple number 1): main-investigator (*IM*) + *JS*. "Variable" couples (couples number 2 - 28) (Table 3). Couple number 1 was coded as being "experienced", because this couple examined more (a total of 45) patients compared with the other "variable" couples (a total of 68 patients examined: 2-3 patients per couple). This design was chosen to evaluate the effect of training and thus a possible increase in reliability of the various selected scales and gadgets when used often (see flowchart).
The examiners received instructions in assessing the various outcome measures. The aim of this training was to strive for a unanimous interpretation regarding the different items of the various scales that were going to be used.

Table 3

Couples of examiners formed by 2 senior neurologists and 6 residents in neurology

Couple numbers	Members	Couple numbers	Members
1	Merkies *(IM)* - Samijn *(JS)*	15	FM – FO
2	IM – van der Meché *(FM)*	16	FM – WM
3	IM - van Doorn *(PD)*	17	FM – RK
4	IM – Opstelten *(FO)*	18	FM – RM
5	IM – Moll *(WM)*	19	PD – FO
6	IM – van Koningsveld *(RK)*	20	PD – WM
7	IM - Meijer *(RM)*	21	PD – RK
8	JS – FM	22	PD – RM
9	JS – PD	23	FO – WM
10	JS - FO	24	FO – RK
11	JS – WM	25	FO – RM
12	JS – RK	26	WM – RK
13	JS – RM	27	WM – RM
14	FM – PD	28	RK – RM

The patients were examined at two different occasions at our outpatient clinic. During the first visit the two members of an appointed pair performed their scores independently and consecutively (usually within 2 hours) (inter-observer measures). Within 2-4 weeks, the patient returned for a second visit and only one investigator of the earlier assigned pair re-examined the patient (intra-observer values) without having access to previous results. The assessment sequence at entry and the examination at the second visit were equally

distributed among the members of an assigned couple. Eventually, each member of a couple examined approximately the same number of patients.

Responsiveness study

Twenty consecutive patients with a sensory-motor GBS (n = 7) or CIDP (n = 13) were longitudinally examined by the same clinician (IM) and all scales were assessed at study entry and 8-13 times in each patient during follow-up. There was a standard follow-up schedule (week 0, 2, 4, 8, 12, 16, 21, 26, 32, 40 and 52) with additional clinical investigations if necessary. At each visit, the patients were requested to judge whether their clinical condition deteriorated (grade 1), remained stable (grade 2) or improved (grade 3) when compared with the last visit ("clinical-judgement-scores"). In each patient, the SF-36 was assessed 5-10 times.

General Notes

To determine the construct convergent validity of the Rydel-Seiffer graduated tuning fork, an electronic device, the Vibrameter (Somedic, Stockholm Sweden, Type III) was also used (150).

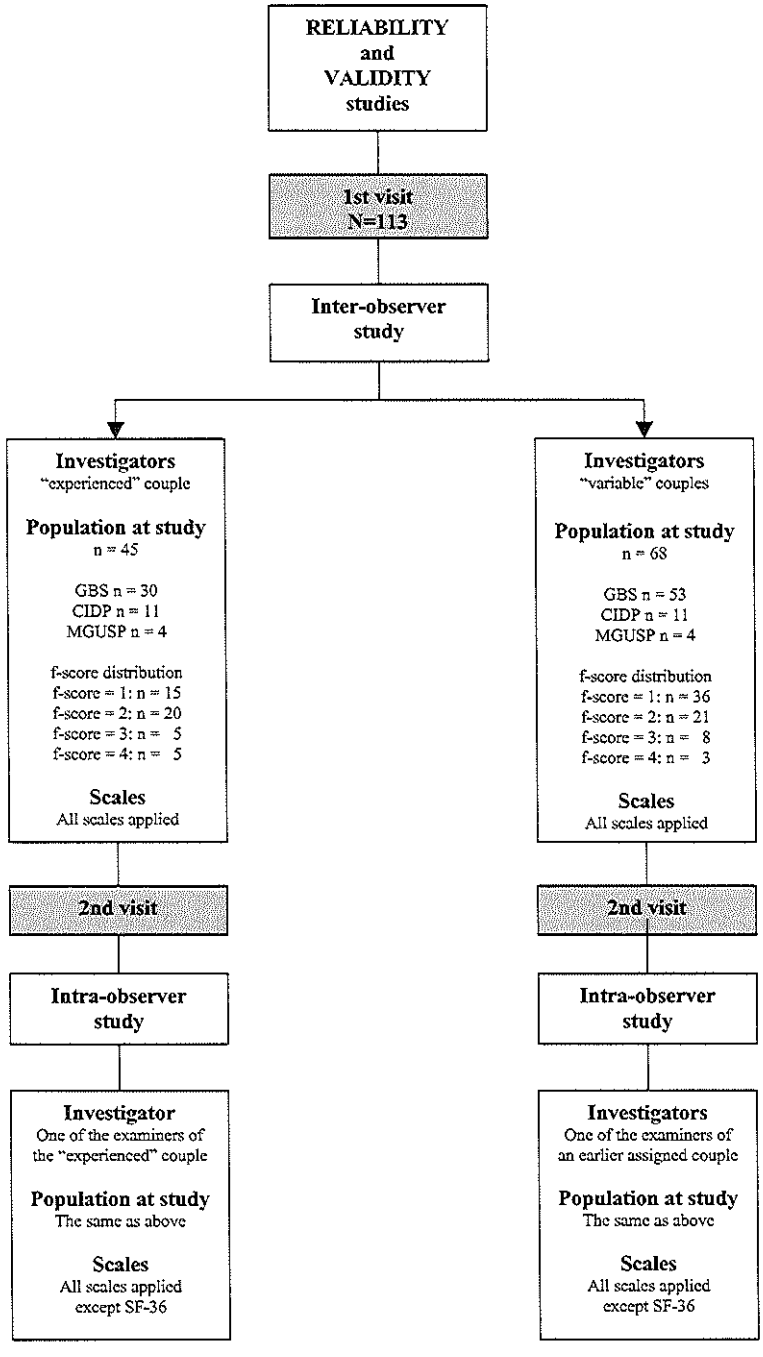
The times to complete the "*INCAT*" sensory sumscore, the Rotterdam 9 Items Handicap scale, and the SF-36 health status were recorded.

In the process of creating the Rotterdam 9 Items Handicap scale, we gathered the opinion of fifty patients with an immune-mediated polyneuropathy (34 GBS, 12, CIDP, 4 MGUSP) regarding the relevance of selected items to create this scale and the clinical applicability of these items. This was done through a telephone interview and by mailing the selected scale items to these patients accompanied by a judgement-form. All suggestions made by these patients were examined and if possible incorporated in the final scale form. Eventually, nine items were selected.

Statistical analyses

Eventually, statistical analyses were performed depending on the postulated objectives and types of data (ordinal or continuous, normally distributed or not, transversal or longitudinally structured). On behalf of the *INCAT* group, a stepwise presentation of the results will be described in the following chapters.

Validity and Reliability study – Flowchart



References

1. Hartung H-P, van der Meché FGA, Pollard JD. Guillain-Barré syndrome, CIDP and other chronic immune-mediated neuropathies. Curr Opin Neurol 1998;11:497-513.

2. Feasby TE. Inflammatory – demyelinating polyneuropathies. Neurol Clin 1992;10:651-670.

3. van der Meché FGA, van Doorn PA. Guillain-Barré syndrome and chronic inflammatory demyelinating polyneuropathy: Immune mechanisms and update on current therapies. Ann Neurol 1995;37(suppl1):S14-S31.

4. Barohn RJ, Saperstein DS. Guillain-Barré syndrome and chronic inflammatory demyelinating polyneuropathy. Semin Neurol 1998;18:49-61.

5. Toyka KV, Hartung H-P. Chronic inflammatory polyneuronitis and related neuropathies. Curr Opin Neurol 1996;9:240-250.

6. Glass JD, Cornblath DR. Chronic inflammatory demyelinating polyneuropathy and paraproteinemic neuropathies. Curr Opin Neurol 1994;7:393-397.

7. Latov N. Pathogenesis and therapy of neuropathies associated with monoclonal gammopathies. Ann Neurol 1995;37(Suppl1):S32-S42.

8. Miescher GC, Steck AJ. Paraproteinaemic neuropathies. Baillière's Clin Neurol 1996;5:219-232.

9. Notermans NC. Monoclonal gammopathy and neuropathy. Curr Opin Neurol 1996;9:334-337.

10. Léger J-M. Multifocal motor neuropathy and chronic inflammatory demyelinating polyradiculoneuropathy. Curr Opin Neurol 1995;8:359-363.

11. Biessels GJ, Franssen H, van den Berg LH, *et al.* Multifocal motor neuropathy. J Neurol 1997;244:143-152.

12. Asbury AK, Arnason BG, Karp HR, *et al.* Criteria for diagnosis of Guillain-Barré syndrome. Ann Neurol 1978;3:565-566.

13. Asbury AK, Cornblath DR. Assessment of current diagnostic criteria for Guillain-Barré syndrome. Ann Neurol 1990;27(suppl):21-24.

14. Ad Hoc Committee WHO-AIREN. Acute onset flaccid paralysis. Geneva: World Health Organisation; 1993.

15. Taly AB, Veerendrakumar M, Das KB, *et al.* Sensory dysfunction in GB syndrome: A clinical and electrophysiological study in 100 patients. Electromyogr Clin Neurophysiol 1997;37:49-52.

16. Dawson DM, Samuels MA, Morris J. Sensory form of acute polyneuritis. Neurology 1988;38:1728-1731.

17. Meulstee J, van der Meché FGA, Dutch Guillain-Barré Study Group. Electrodiagnostic criteria for polyneuropathy and demyelination: application in 135 patients with Guillain-Barré syndrome. J Neurol Neurosurg Psychiatry 1995;59:482-486.

18. Ad Hoc Subcommittee of the American Academy of Neurology AIDS Task Force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

19. Hughes RAC, Sanders E, Hall S, *et al.* Subacute idiopathic demyelinating polyradiculoneuropathy. Arch Neurol 1992;49:612-616.

20. Ropper AH, Wijdicks EFM, Truax BT. Guillain-Barré syndrome. Contemporary neurology series. Philadelphia: FA Davis Company, 1991.

21. Gibbels E, Giebisch U. Natural course of acute and chronic monophasic inflammatory demyelinating polyneuropathies (IDP). Acta Neurol Scand 1991;85:282-291.

22. Dyck PJ, Prineas J, Pollard JD. Chronic inflammatory demyelinating polyneuropathy. In: Dyck PJ, Thomas PK, eds. Peripheral neuropathy. 3rd ed. Philadelphia: Saunders, 1993: 1498-1518.

23. Gorson KC, Allam G, Ropper AH. Chronic inflammatory demyelinating polyradiculoneuropathy: clinical features and response to treatment in 67 consecutive patients with and without a monoclonal gammopathy. Neurology 1997;48:321-328.

24. Oh SJ, Joy JL, Kuruoglu R. "Chronic sensory demyelinating neuropathy": chronic inflammatory demyelinating polyneuropathy presenting as a pure sensory neuropathy. J Neurol Neurosurg Psychiatry 1992;55:677-680.

25. Franssen H, Vermeulen M, Jennekens FGI. Chronic inflammatory neuropathies. In: Emery AEH, ed. Diagnostic criteria for neuromuscular disorders. London, Royal Society of Medicine Press, 1997: 53-59.

26. Dyck PJ, Lais AC, Ohta M, *et al.* Chronic inflammatory polyradiculoneuropathy. Mayo Clin Proc 1975;50:621-637.

27. Steck AJ. Inflammatory neuropathy: pathogenesis and clinical features. Curr Opin Neurol Neurosurg 1992;5:633-637.

28. Bromberg MB, Feldman EL, Albers JW. Chronic inflammatory demyelinating polyradiculoneuropathy: comparison of patients with and without an associated monoclonal gammopathy. Neurology 1992;42:1157-1163.

29. Simmons Z, Albers JW, Bromberg MB, *et al.* Presentation and initial clinical course in patients with chronic inflammatory demyelinating polyradiculoneuropathy: Comparison of patients without and with monoclonal gammopathy. Neurology 1993;43:2202-2209.

30. Simmons Z, Albers JW, Bromberg MB, *et al.* Long-term follow-up of patients with chronic inflammatory demyelinating polyradiculoneuropathy, without and with monoclonal gammopathy. Brain 1995;118:359-368.

31. Pestronk A, Cornblath DR, Ilyas AA, *et al.* A treatable multifocal motor neuropathy with antibodies to GM1 ganglioside. Ann Neurol 1988;24:73-78.

32. Parry GJ. Motor neuropathy with multifocal conduction block. In: Dyck PJ and Thomas PK. Peripheral neuropathy. Saunders, Philadelphia. 3rd ed. 1993; 1518-1524.

33. Kornberg AJ and Pestronk A. Chronic motor neuropathies: Diagnosis, Therapy, and Pathogenesis. Ann Neurol 1995;37(suppl 1):S43-S50.

34. Oh SJ, Claussen GC, Odabasi Z, *et al.* Multifocal demyelinating motor neuropathy: Pathologic evidence of 'inflammatory demyelinating polyradiculoneuropathy'. Neurology 1995;45:1828-1832.

35. Krendel DA. Biopsy findings link multifocal motor neuropathy to chronic inflammatory demyelinating polyneuropathy. Ann Neurol 1996;40:948-949.

36. Sharoqi IA, Sadoh DR, Koutroumanidis M, *et al.* Chronic relapsing multifocal sensory-motor neuropathy with conduction block. J Peripheral Nervous System 1998;3:133-136.

37. Lewis RA, Sunner AJ, Brown MJ, *et al.* Multifocal demyelinating neuropathy with persistent conduction block. Neurology 1982;32:958-964.

38. World Health Organisation. International classification of impairments, disabilities, and handicaps. Geneva: WHO, 1980.

39. Aaronson NK. Quality of life: What is it? How should it be measured? Oncology 1988;2:69-76.

40. Wade DT. Measurements in neurological rehabilitation. Oxford University press; Oxford: 1992.

41. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. J Neurol Neurosurg Psychiatry 1996;60:127-130.

42. Feinstein AR. Clinimetrics. Yale University Press; New Haven and London, 1987.

43. Liang MH. Evaluating measurement responsiveness. J Rheumatol 1995;22:1191-1192.

44. Fortin PR, Stucki G, Katz JN. Measuring relevant change: an emerging challenge in rheumatologic clinical trials. Arthritis Rheum 1995;38:1027-1030.

45. Hughes RAC, Newsom-Davis JM, Perkin GD, *et al.* Controlled trial prednisolone in acute polyneuropathy. Lancet 1978;2:750-753.

46. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. New York, Oxford University Press, 2nd ed., 1998.

47. Nunnally JC. Psychometric theory. New York: McGraw Hill, 1978.

48. Medical Research Council. Aids to the investigation of the peripheral nervous system. London: Her Majesty's Stationary Office, 1943:1-2.

49. Kleyweg RP, van der Meché FGA, Schmitz PIM. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré syndrome. Muscle Nerve 1991;14:1103-1109.

50. Dyck PJ, Kratz KM, Lehman KA, *et al.* The Rochester diabetic neuropathy study: Design, criteria for types of neuropathy, selection bias, and reproducibility of neuropathic tests. Neurology 1991;41:799-807.

51. Dyck PJ, Karnes JL, O'Brien PC, *et al.* The Rochester diabetic neuropathy study: Reassessment of tests and criteria for diagnosis and staged severity. Neurology 1992;42:1164-1170.

52. van Swieten JC, Koudstaal PJ, Visser MC, *et al.* Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988;19:604-607.

53. Molenaar DS, Vermeulen M, de Visser M, *et al.* Impact of neurologic signs and symptoms on functional status in peripheral neuropathies. Neurology 1999;52:151-156.

54. Jacobs HM, Lutik A, Touw-Otten FWMM, *et al.* The sickness impact profile; results of a validation study of the Dutch version. Ned Tijdschr Geneeskd 1990;134:1950-1954.

55. The French cooperative group on plasma exchange in Guillain-Barré syndrome. Appropriate number of plasma exchanges in Guillain-Barré syndrome. Ann Neurol 1997;41:298-306.

56. Plasma exchange/Sandoglobulin Guillain-Barré syndrome trial group. Randomised trial of plasma exchange, intravenous immunoglobulin, and combined treatments in Guillain-Barré syndrome. Lancet 1997;349:225-230.

57. Van der Meché FGA, Schmitz PIM, and the Dutch Guillain-Barré study group. A randomised trial comparing intravenous immune globulin and plasma exchange in Guillain-Barré syndrome. N Engl J Med 1992;326:1123-1129.

58. Winer JB, Hughes RAC, Osmond C. A prospective study of acute idiopathic neuropathy. 1. Clinical features and their prognostic value. J Neurol Neurosurg Psychiatry 1988;51:605-612.

59. Bernsen RAJAM, Jacobs HM, de Jager AEJ, *et al.* Residual health status after Guillain-Barré syndrome. J Neurol Neurosurg Psychiatry 1997;62:637-640.

60. The Dutch Guillain-Barré study group. Treatment of Guillain-Barré syndrome with high-dose immune globulins combined with methylprednisolone: A pilot study. Ann Neurol 1994;35:749-752.

61. McKhann GM, Cornblath DR, Ho T, *et al.* Clinical and electrophysiological aspects of acute paralytic disease of children and young adults in northern China. Lancet 1991;338:593-597.

62. Gregson NA, Koblar S, Hughes RAC. Antibodies to gangliosides in Guillain-Barré syndrome: specificity and relationship to clinical features. Q J Med 1993;86:111-117.

63. Hung KL, Wang HS, Liou WY, *et al.* Guillain-Barré syndrome in children: a cooperative study in Taiwan. Brain Dev 1994;16:204-208.

64. Lyu RK, Tang LM, Cheng SY, *et al.* Guillain-Barré syndrome in Taiwan: a clinical study of 167 patients. J Neurol Neurosurg Psychiatry 1997;63:494-500.

65. Sedano MJ, Calleja J, Canga E, *et al.* Guillain-Barré syndrome in Cantabria, Spain. An epidemiological and clinical study. Acta Neurol Scand 1994;89:287-292.

66. Howlett WP, Vedeler CA, Nyland H, *et al.* Guillain-Barré syndrome in northern Tanzania: a comparison of epidemiological and clinical findings with western Norway. Acta Neurol Scand 1996;93:44-49.

67. Visser LH, van der Meché FGA, Meulstee J, *et al.* Cytomegalovirus infection and Guillain-Barré syndrome: the clinical, electrophysiologic, and prognostic features. Neurology 1996;47:668-673.

68. Emilia-Romagna Study Group on Clinical and Epidemiological Problems in Neurology. Guillain-Barré syndrome variants in Emilia-Romagna, Italy, 1992-3:incidence, clinical features, and prognosis. J Neurol Neurosurg Psychiatry 1998;65:218-224.

69. Gupta SK, Taly AB, Suresh TG, *et al.* Acute idiopathic axonal neuropathy (AIAN): a clinical and electrophysiological observation. Acta Neurol Scand 1994;89:220-224.

70. Visser LH, van der Meché FGA, Meulstee, *et al.* Risk factors for treatment related clinical fluctuations in Guillain-Barré syndrome. J Neurol Neurosurg Psychiatry 1998;64:242-244.

71. de Jager AEJ, Minderhoud JM. Residual signs in severe Guillain-Barré syndrome: analysis of 57 patients. J Neurol Sci 1991;104:151-156.

72. de Jager AEJ, Sluiter HJ. Clinical signs in severe Guillain-Barré syndrome: analysis of 63 patients. J Neurol Sci 1991;104:143-150.

73. Grand'Maison F, Feasby TE, Hahn AF, *et al.* Recurrent Guillain-Barré syndrome. Clinical and laboratory features. Brain 1992;115:1093-1106.

74. Bril V, Ilse WK, Pearce R, *et al.* Pilot trial of immunglobulin versus plasma exchange in patients with Guillain-Barré syndrome. Neurology 1996;46:100-103.

75. Epstein MA, Sladky JT. The role of plasmapheresis in childhood Guillain-Barré syndrome. Ann Neurol 1990;28:65-69.

76. Rees JH, Soudain SE, Gregson NA, *et al.* Campylobacter jejuni infection and Guillain-Barré syndrome. N Engl J Med 1995;333:1374-1379.

77. Castro LHM, Ropper AH. Human immune globulin infusion in Guillain-Barré syndrome: Worsening during and after treatment. Neurology 1993;43:1034-1036.

78. Guillain-Barré syndrome Steroid Trial group. Double-blind of intravenous methylprednisolone in Guillain-Barré syndrome. Lancet 1993;341:586-590.

79. Taly AB, Gupta SK, Vasanth A, *et al.* Critically ill Guillain-Barré syndrome. J Assoc Physicians India 1994;42:871-874.

80. Baba M, Matsunaga M, Narita S, *et al.* Recurrent Guillain-Barré syndrome in Japan. Intern Med 1995;34:1015-1018.

81. Irie S, Saito T, Kanazawa N, *et al.* Relationships between anti-ganglioside antibodies and clinical characteristics of Guillain-Barré syndrome. Intern Med 1997;36:607-612.

82. Ropper AH, Albers JW, Addison R. Limited relapse in Guillain-Barré syndrome after plasma exchange. Arch Neurol 1988;45:314-315.

83. Bradshaw DY, Jones HR. Guillain-Barré syndrome in children: clinical course, electrodiagnosis, and prognosis. Muscle Nerve 1992;15:500-506.

84. Kleyweg RP, van der Meché FGA, Loonen MCB, *et al.* The natural history of the Guillain-Barré syndrome in 18 children and 50 adults. J Neurol Neurosurg Psychiatry 1989;52:853-856.

85. Haaβ A, Trabert W, Greβnich N, *et al.* High-dose steroid therapy in Guillain-Barré syndrome. J Neuroimmunol 1988;20:305-308.

86. Smith GDP, Hughes RAC. Plasma exchange treatment and prognosis of Guillain-Barré syndrome. Q J Med 1992;85:751-760.

87. Bahemuka M. Guillain-Barré syndrome in Kenya: a clinical review of 54 patients. J Neurol 1988;235:418-421.

88. Vedeler CA, Wik E, Nyland H. The long-term prognosis of Guillain-Barré syndrome. Evaluation of prognostic factors including plasma exchange. Acta Neurol Scand 1997;95:298-302.

89. Al-Din ASN, Anderson M, Eeg-Olofsson O, *et al.* Neuro-ophthalmic manifestations of the syndrome of ophthalmoplegia, ataxia and areflexia. Acta Neurol Scand 1994;89:87-94.

90. Gorson KC, Ropper AH, Clark BD, *et al.* Treatment of chronic inflammatory demyelinating polyneuropathy with interferon-α 2a. Neurology 1998;50:84-87.

91. Choudhary PP, Hughes RAC. Long-term treatment of chronic inflammatory demyelinating polyradiculoneuropathy with plasma exchange or intravenous immunoglobulin. Q J Med 1995;88:493-502.

92. Nevo Y, Pestronk A, Kornberg AJ, *et al.* Childhood chronic inflammatory demyelinating neuropathies: Clinical course and long-term follow-up. Neurology 1996;47:98-102.

93. Simmons Z, Wald JJ, Albers JW. Chronic inflammatory demyelinating polyradiculoneuropathy in children: I. Presentation, electrodiagnostic studies, and initial clinical course, with comparison to adults. Muscle Nerve 1997;20:1008-1015.

94. Simmons Z, Wald JJ, Albers JW. Chronic inflammatory demyelinating polyradiculoneuropathy in children: II. Long-term follow-up, with comparison to adults. Muscle Nerve 1997;20:1569-1575.

95. Barohn RJ, Kissel JT, Warmolts JR, *et al.* Chronic inflammatory demyelinating polyradiculoneuropathy. Clinical characteristics, course, and recommendations for diagnostic criteria. Arch Neurol 1989;46:878-884.

96. Van Doorn PA, Vermeulen M, Brand A, *et al.* Intravenous immunoglobulin treatment in patients with chronic inflammatory demyelinating polyradiculoneuropathy. Clinical and laboratory characteristics associated with improvement. Arch Neurol 1991;48:217-220.

97. Mariette X, Chastang C, Clavelou P, *et al.* A randomised clinical trial comparing interferon-α and intravenous immunoglobulin in polyneuropathy associated with monoclonal IgM. J Neurol Neurosurg Psychiatry 1997;63:28-34.

98. Wertman E, Argov Z, Abrmasky O. Chronic inflammatory demyelinating polyradiculoneuropathy: Features and prognostic factors with corticosteroid therapy. Eur Neurol 1988;28:199-204.

99. Hahn AF, Bolton CF, Zochodne D, *et al.* Intravenous immunoglobulin treatment in chronic inflammatory demyelinating polyneuropathy. A double-blind, placebo-controlled, crossover study. Brain 1996;119:1067-1077.

100. Hahn AF, Bolton CF, Pillay N, *et al.* Plasma-exchange therapy in chronic inflammatory demyelinating polyneuropathy. A double-blind, sham-controlled, crossover study. Brain 1996;119:1055-1066.

101. Dyck PJ, Low PA, Windebank AJ, *et al.* Plasma exchange in polyneuropathy associated with monoclonal gammopathy of undetermined significance. N Engl J Med 1991:325:1482-1486.

102. Suarez GA, Kelly JJ. Polyneuropathy associated with monoclonal gammopathy of undetermined significance: Further evidence that IgM-MGUS neuropathies are different than IgG-MGUS. Neurology 1993;43:1304-1308.

103. Smith IS. The natural history of chronic demyelinating neuropathy associated with benign IgM paraproteinaemia. A clinical and neurophysiological study. Brain 1994;117:949-957.

104. Cornblath DR, Chaudhry V, Griffin JW. Treatment of chronic inflammatory demyelinating polyneuropathy with intravenous immunoglobulin. Ann Neurol 1991;30:104-106.

105. Bromberg MB, Feldman EL, Jaradeh S, *et al.* Prognosis in long-term immunosuppressive treatment of refractory chronic inflammatory demyelinating polyradiculoneuropathy. J Clin Epidemiol 1992;45:47-52.

106. French cooperative group on plasma exchange in Guillain-Barré syndrome. Plasma exchange in Guillain-Barré syndrome: One-year follow-up. Ann Neurol 1992;32:94-97.

107. Kleyweg RP, van der Meché FGA. Treatment related fluctuations in Guillain-Barré syndrome after high-dose immunoglobulins or plasma-exchange. J Neurol Neurosurg Psychiatry 1991;54:957-960.

108. Good JL, Chehrenama M, Mayer RF, *et al.* Pulse cyclophosphamide therapy in chronic inflammatory demyelinating polyneuropathy. Neurology 1998;51:1735-1738.

109. Molenaar DSM, van Doorn PA, Vermeulen M. Pulsed high dose dexamethasone treatment in chronic inflammatory demyelinating polyneuropathy: a pilot study. J Neurol Neurosurg Psychiatry 1997;62:388-390.

110. Vermeulen M, van Doorn PA, Brand A, *et al.* Intravenous immunoglobulin treatment in patients with chronic inflammatory demyelinating polyneuropathy: a double blind, placebo controlled study. J Neurol Neurosurg Psychiatry 1993;56:36-39.

111. Hattori N, Ichimura M, Aoki S, *et al.* Clinicopathological features of chronic inflammatory demyelinating polyradiculoneuropathy in childhood. J Neurol Sci 1998;154:66-71.

112. Maisonobe T, Chassande B, Vérin M, *et al.* Chronic dysimmune demyelinating polyneuropathy: a clinical and electrophysiological study of 93 patients. J Neurol Neurosurg Psychiatry 1996;61:36-42.

113. Krarup C, Trojaborg W. Sensory pathophysiology in chronic acquired demyelinating neuropathy. Brain 1996;119:257-270.

114. Dyck PJ, Litchy WJ, Kratz KM, *et al.* A plasma exchange versus immune globulin infusion trial in chronic inflammatory demyelinating polyradiculoneuropathy. Annals Neurol 1994;36:838-845.

115. Gorson KC, Ropper AH. Axonal neuropathy associated with monoclonal gammopathy of undetermined significance. J Neurol Neurosurg Psychiatry 1997;63:163-168.

116. Nobile-Orazio E, Francomano E, Daverio R, *et al.* Anti-myelin-associated glycoprotein IgM antibody titers in neuropathy associated with macroglobulinemia. Ann Neurol 1989;26:543-550.

117. Notermans NC, Lokhorst HM, Franssen H, *et al.* Intermittent cyclophosphamide and prednisone treatment of polyneuropathy associated with monoclonal gammopathy of undetermined significance. Neurology 1996;47:1227-1233.

118. Gosselin S, Kyle RA, Dyck PJ. Neuropathy associated with monoclonal gammopathies of undetermined significance. Ann Neurol 1991;30:54-61.

119. Oksenhendler E, Chevret S, Léger J-M, *et al.* Plasma exchange and chlorambucil in polyneuropathy associated with monoclonal IgM gammopathy. J Neurol Neurosurg Psychiatry 1995;59:243-247.

120. Nobile-Orazio E, Manfredini E, Carpo M, *et al.* Frequency and clinical correlates of anti-neural IgM antibodies in neuropathy associated with IgM monoclonal gammopathy. Ann Neurol 1994;36:416-424.

121. Notermans NC, Wokke JHJ, Lokhorst HM, *et al.* Polyneuropathy associated with monoclonal gammopathy of undetermined significance. A prospective study of the prognostic value of clinical and laboratory abnormalities. Brain 1994;117:1385-1393.

122. Chassande B, Léger JM, Younes-Chennoufi AB, *et al.* Peripheral neuropathy associated with IgM monoclonal gammopathy: correlations between M-protein antibody activity and clinical/electrophysiological features in 40 cases. Muscle Nerve 1998;21:55-62.

123. Dalakas MC, Quarles RH, Farrer RG, *et al.* A controlled study of intravenous immunoglobulin in demyelinating neuropathy with IgM gammopathy. Ann Neurol 1996;40:792-795.

124. Ellie E, Vital A, Steck A, *et al.* Neuropathy associated with "benign" anti-myelin-associated glycoprotein IgM gammopathy: clinical, immunological, neurophysiological, pathological findings and response to treatment in 33 cases. J Neurol 1996;243:34-43.

125. Yeung KB, Thomas PK, King RHM, *et al.* The clinical spectrum of peripheral neuropathies associated with benign monoclonal IgM, IgG and IgA paraproteinaemia. Comparative clinical, immunological and nerve biopsy findings. J Neurol 1991;238:383-391.

126. Nobile-Orazio E, Barbieri S, Baldini L, *et al.* Peripheral neuropathy in monoclonal gammopathy of undetermined significance: prevalence and immunopathogenetic studies. Acta Neurol Scand 1992;85:383-390.

127. Vrethem M, Cruz M, Wen-Xin H, *et al.* Clinical, neurophysiological and immunological evidence or polyneuropathy in patients with monoclonal gammopathies. J Neurol Sci 1993;114:193-199.

128. Miralles GD, O'Fallon JR, Talley NJ. Plasma-cell dyscrasia with polyneuropathy. The spectrum of POEMS syndrome. N Engl J Med 1992;327:1919-1923.

129. Pestronk A, Griffin J, Feldman EL, *et al.* Polyneuropathy syndromes associated with serum antibodies to sulfatide and myelin-associated glycoprotein. Neurology 1991;41:357-362.

130. Kelly JJ, Adelman LS, Berkman E, *et al.* Polyneuropathies associated with IgM monoclonal gammopathies. Arch Neurol 1988:45:1355-1359.

131. Lennon SM, Koblar S, Hughes RAC, *et al.* Reasons for persistent disability in Guillain-Barré syndrome. Clin Rehabil 1993;7:1-8.

132. Stewart G, Kidd D, Thompson AJ. The assessment of handicap: an evaluation of the Environmental Status Scale. Disabil Rehabil 1995;17:312-316.

133. Bombardier C, Tugwell P. A methodological framework to develop and select indices for clinical trials: Statistical and judgmental approaches. J Rheumatol 1982;9:753-757.

134. Rydel A, Seiffer W. Untersuchungen über das vibrationsgefühl oder die sog. "Knochensensibilität" (Pallästhesie). Archiv für Psychiatrie und Nervenkrankenheiten 1903;37:488-536.

135. Fünfgeld EW. Vigorimeter: Zur Kraftmessung der Hand und zur Simulationsprüfung. Dtsch Med Wschr 1966;49:2214-2216.

136. Krupp LB, LaRocca NG, Muir-Nash J, *et al.* The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 1989;46:1121-1123.

137. Jones E, Hanly JG, Mooney R, *et al.* Strength and function in the normal and rheumatoid hand. J Rheumatol 1991;18:1313-1318.

138. Desrosiers J, Hébert R, Bravo G, *et al.* Comparison of the Jamar Dynamometer and the Martin Vigorimeter for grip strength measurements in a healthy elderly population. Scand J Rehab Med 1995;27:137-143.

139. Link L, Lukens S, Bush MA. Spherical grip strength in children 3 to 6 years of age. Amer J Occup Ther 1995;49:318-326.

140. Robertson A, Deitz J. A description of grip strength in preschool children. Amer J Occup Ther 1988;42:647-652.

141. Claus D, Carvalho VP, Neundörfer B, *et al.* Zur unterzuchung des vibrationsempfindens. Nervenarzt 1988;59:138-142.

142. Crausaz FM, Clavel S, Liniger C, *et al.* Additional factors associated with plantar ulcers in diabetic neuropathy. Diab Med 1988;5:771-775.

143. Liniger C, Albeanu A, Bloise D, *et al.* The tuning fork revisited. Diab Care 1990;7:859-864.

144. Thivolet C, El Farkh J, Petiot A, *et al.* Measuring vibration sensation with graduated tuning fork. Diab Care 1990;13:1077-1080.

145. Sharrack B, Hughes RAC, Soudain S. Guy's neurological disability scale [abstract]. Neurology 1996;243(suppl 2):S32.

146. Ware Jr. JE, Snow KK, Kosinski M, *et al.* SF-36 health survey. Manual and interpretation guide. Boston: The Health Institute, New England Medical Center, 1997.

147. Ware Jr. JE, Kosinski M, Keller SD. SF-36 physical and mental health summary scales: a user's manual. Boston: The Health Assessment Lab, New England Medical Center, 1994.

148. Aaronson NK, Muller M, Cohen PDA, *et al.* Translation, validation, and norming of the Dutch language version of the SF-36 health survey in community and chronic disease populations. J Clin Epidemiol 1998;51:1055-1068.

149. Ware Jr. JE, Gandek B, Kosinski M, *et al.* The equivalence of SF-36 summary health scores estimated using standard and country-specific algorithms in 10 countries: results form the IQOLA project. J Clin Epidemiol 1998;51:1167-1170.

150. Goldberg JM, Lindblom U. Standardised method of determining vibratory perception thresholds for diagnosing and screening in neurological investigation. J Neurol Neurosurg Psychiatry 1979;42:793-803.

Chapter 2

Measuring vibration threshold with a graduated tuning fork in normal ageing and patients with polyneuropathy

Merkies ISJ[1], van Koningsveld R[1], Schmitz PIM[2], van der Meché FGA[1], van Doorn PA[1], for the Inflammatory Neuropathy Cause And Treatment (INCAT) group[3]

From the Department of [1]Neurology, [2]Department of Statistics - Daniel den Hoed Cancer Centre, Erasmus University Medical Centre Rotterdam, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands.

[3]Members of the INCAT-group: Aubry J, PhD; Baumann N, PhD; Bouche P, MD, PhD; Comi G, MD, PhD; Dalakas MC, MD, PhD; Hartung HP, MD, PhD; Hughes RAC, MD, PhD; Illa I, MD, PhD; Knapp MRJ, PhD; Léger J-M, MD, PhD; Nemni R, MD, PhD; Nobile-Orazio E, MD, PhD; Toyka KV, MD, PhD; van de Bergh P, MD, PhD; van der Meché FGA, MD, PhD; van Doorn PA, MD, PhD; Willison HJ, MD, PhD; Younes-Chennoufi B, PhD

J Neurol Neurosurg Psychiatry 1998;65:743-747; published with permission from the BMJ Publishing Group

published under the former name Martina ISJ

Abstract

Objective: To provide clinical useful vibration threshold normal values.
Methods: The graduated Rydel-Seiffer tuning fork was evaluated in 198 healthy controls and 59 patients with a polyneuropathy. The measures were done in triplicate at 4 locations: the distal interphalangeal joint of the index finger, ulnar styloid process, interphalangeal joint of the hallux, and internal malleolus. The values obtained with this tuning fork in healthy controls and patients with polyneuropathy were compared with the values of an electronic device, the Vibrameter.
Results: Vibration sense was better perceived in the arms compared with the legs. There was a significant age related decline of vibration sense at all locations. The values from the Rydel-Seiffer tuning fork and the Vibrameter were significantly correlated in both groups. The sensitivity of these two instruments for the 4 sites examined in the polyneuropathy group ranged from 29-76% and 31-73%, respectively and was the highest at the hallux for both instruments.
Conclusion: This study provides clinical useful normal values of vibration threshold for the Rydel-Seiffer tuning fork. This is a simple and easily applicable instrument that assesses vibration sense semi-quantitatively and should therefore have a place in the routine neurological examination.

Introduction

Abnormalities in sensory qualities are frequent complaints of patients with peripheral neuropathies. Impairment of vibration sense, especially starting distally in the limbs, may be a sign of peripheral nerve dysfunction (1). Traditionally, a non-graduated tuning fork, which was invented in 1711 and introduced by Bonnafont as a diagnostic test in general medicine, is used for the evaluation of the vibration sense (2). This simple instrument measures the presence or absence of vibration, but unfortunately does not quantitatively provide the degree of dysfunction of vibration sense. It is of clinical importance that the vibration sense should be measured quantitatively and consistently. For this purpose several electronic devices have been developed, such as the Biothesiometer, the Optacon, and the Computer-assisted-sensory-examination (3-5). Although these instruments are useful in experimental studies, they are of little use in daily practice. The size of these apparatus, the duration of examination, the demand on patients' co-operation, and their cost are in sharp contrast with the needs of simple and valid instruments for use in the routine neurological investigation. With the introduction of the 64 Hz graduated tuning fork by Rydel and Seiffer (Martin, Tuttlingen, Germany) in 1903 it seemed as if all the needs were fulfilled in providing an instrument which was easy to apply, inexpensive, and reliable for quantifying impairment of vibration sense (6). However, this tool has been largely neglected by neurologists and thus far, only a few papers have reported on its use (table 1, page 45; 7-14). Clinically useful normal values for vibration sense using this tuning fork were provided by one study only (7). These values were obtained by examining 73 healthy controls only at the hallux, which hampers their general applicability. Other studies acquired their own normal vibration values, but unfortunately from only a few healthy controls (8,11). No study has provided normal values for the arms. The primary aim of this study was to obtain specific vibration threshold normal values for the Rydel-Seiffer tuning fork at 4 examination sites in a large number of healthy controls. In addition, the values achieved using this device in healthy controls and in patients with a polyneuropathy were compared with those obtained with an electronic device, the Vibrameter (15).

Participants and Methods

Healthy controls

Healthy controls were recruited from hospital personnel, companions (relatives, friends) of patients visiting our outpatient clinic, and from homes for elderly people. Two hundred and fourteen potential controls were interviewed and examined. Those with sensory symptoms, sensory signs including absent vibration sense, a history of alcohol abuse, using drugs which may cause a polyneuropathy or influence their cooperation, and those with a disease which might induce a polyneuropathy, were excluded from the study. The selected participants had a lucid consciousness and their history did not show any mental or psychological illness. Standard neurological examination was performed with special interest in the sensory qualities. Six age groups (<40, 40-49, 50-59, 60-69, 70-79, ≥80 years) were formed. One hundred and ninety eight healthy controls (93 men; 105 women; mean age 55.1 (SD 18.0) years; range 19-93) stratified for sex, were enrolled in this study. Each age group consisted of approximately 35 participants. Sixteen individuals (9 men; 7 women; mean age 81 (SD

8.5) years; range 65-93 years) were excluded from analysis, based on absent vibration perception in at least one region tested. Eight had an absent vibration sense at the hallux and the internal malleolus, seven only at the hallux and one only at the internal malleolus. Nine subjects had absent and four decreased (-3/-3) ankle jerks. The ankle reflexes were normal in the remaining 3 persons. These patients were considered to have subclinical disease. The problem of the selection of participants for normal vibration threshold values is discussed later.

Patients with polyneuropathy

Fifty-nine ambulatory patients from our outpatient clinic (31 men; 28 women; mean age 56.8 (SD 15.9) years; range 14-87 years) with a clinical and electrophysiologically supported polyneuropathy were enrolled in this study. They were stratified into age groups, as described above, and sex. Each age group consisted of approximately 10 participants. These patients had a mild polyneuropathy with limited sensory disturbances and walking problems. The aetiologies of the polyneuropathies were diabetes mellitus (11 cases), hereditary motor sensory neuropathy type I (5 cases) and type II (2 cases), systemic disease (3 cases), amyloidosis (3 cases), drug induced (2 cases), thyroid dysfunction (1 case), and vitamin B12 deficiency (1 case). Fourteen patients had a chronic idiopathic axonal polyneuropathy. In 17 cases no cause was determined.

Assessment tools

The Rydel-Seiffer tuning fork is a graduated fork, which determines the ability of individuals to discriminate various vibration intensities. The two arms of this tuning fork bear calibrated weights at their extremities (figure 1, page 44; see also appendix I, page 207) (6,8). A triangle and an arbitrary scale from 0 (minimum score) to 8 (maximum score) imprinted on the weights allow assessment of vibration threshold. Once the arms are swinging, the fork vibrates at 64 Hz and the triangles on the weights appear double. The intersection of these two virtual triangles moves from 0 to 8 in an exponential way with decreasing vibration amplitude of the arms. The vibration extinction threshold is considered as the nearest value to the apparent point of intersection of the virtual triangles when the subject indicates that vibration is no longer perceived. The Vibrameter (Somedic, Stockholm, Sweden, Type III) is a device that determines vibration sense electronically. It has been extensively described by Goldberg and Lindblom and consists essentially of a hand held vibrating probe that vibrates at 100 Hz (15). The vibration amplitude of this probe increases from zero to a maximum of 399.9 μm.

Tests procedures

All participants gave informed consent before the study. One investigator (RvK) performed all measurements and most assessments took place in the morning. Vibration threshold was assessed at the dorsum of the distal interphalangeal joint of the index finger, ulnar styloid process, dorsum of the interphalangeal joint of the hallux, and internal malleolus. The examination was always performed at the right side of the body with the exception of those with an injury or malformation at that particular side. The measures were obtained in a quiet, comfortably warm, central heating temperature controlled room at our outpatient clinic or in bedrooms of homes for elderly people (20°-22° C). The tuning fork was applied as perpendicular as possible resting on its own weight with the arms of the fork swinging

maximally. The participants were asked to lie at ease in supine position and indicate the moment when they no longer perceived the decreasing vibration stimulus. The readings of three repeated tests were averaged and considered the vibration threshold for that particular site of examination. In addition, vibration sense was assessed with the Vibrameter in half of the selected healthy controls of each age group, stratified for sex (48 men; 52 women; total n = 100), and all patients with polyneuropathy. Vibration threshold was measured according to the method of limits (15). The subjects were asked to indicate when the vibration stimulus was felt for the first time (perception threshold) and when this stimulus disappeared again (disappearance threshold). The average of these 3-paired measurements was considered the vibration threshold at the location investigated. The Vibrameter was applied resting on its own weight and its equilibration was electronically controlled. Each cycle of measurement included catch trials with a resting probe.

Statistics

Vibration threshold reference values for the Rydel-Seiffer tuning fork were calculated for the four sites examined, depending on age and sex, using linear and quadratic regression analysis at a chosen specificity of 95%. The 5% lower limits for the Rydel-Seiffer tuning fork were estimated for each site of examination as the mean - 1.65*Sres. The mean was defined as the mean vibration threshold for a certain age (and gender) and Sres was considered the residual standard deviation around the regression line. The 95th percentile vibration threshold values for the Vibrameter were also calculated. The obtained limits were estimated and further used to determine new specificities and sensitivities for both instruments. A vibration sense was considered to be abnormal if the corresponding value was below the 5% lower limit when examined with the tuning fork or above the 95th percentile vibration value when investigated with the Vibrameter at the same site of examination. The correlation between the two instruments was analysed by Spearman's rank correlation coefficient. All analyses were performed using Stata 5.0 for Windows 95 (Stata Statistical Software: Release 5.0. 702 University Drive East, College station, TX: Stata Corporation; 1997). A p-value < 0.05 was considered to be significant.

Results

The whole procedure of assessing vibration sense with the Rydel-Seiffer tuning fork took about 5 minutes, whereas measurements with the Vibrameter needed 10-15 minutes to be completed. In the healthy controls, there was a significant regression between age and the acquired vibration threshold values for the Rydel-Seiffer tuning fork at each site of examination ($p < 0.0001$). The corresponding graphics show for each location that the vibration threshold values decrease with ageing (figures 2-5, page 46-47). The obtained equations were linear at each site examined, except for the internal malleolus. At the latter, the vibration thresholds had a quadratic regression on age. Only at this location, a significant regression was also found between the vibration values and sex ($p < 0.0001$). The calculated vibration thresholds at this side were 0.51 higher (scale 0-8) for women compared to men. The 5% lower limit values were calculated for each age (and sex) at each location and additionally translated for use in clinical practice (table 2, page 48). The values for vibration sense were higher in the arms than in the legs. In addition, there was a significant ($p <$

0.0001) negative correlation between the results of the two instruments at all sites examined. In the healthy controls, the Spearman's rank correlation coefficient (Srcc) ranged from -0.46 at the index finger to a maximum of -0.65 at the hallux. In the group of patients with polyneuropathy, the Srcc showed almost the same pattern, ranging from -0.46 at the internal malleolus to a maximum of -0.71 at the hallux.

When studying the patients with a mild polyneuropathy at a chosen specificity of 95%, the corresponding sensitivity of the Rydel-Seiffer tuning fork ranged from 29% at the ulnar styloid process to a maximum of 76% at the hallux. The Vibrameter had a sensitivity that ranged from 31% at the index finger to a maximum of 73% at the hallux (table 3, page 48). The Rydel-Seiffer tuning fork detected at the locations in the legs a total of 8 patients more compared to the Vibrameter. There was no difference in sensitivity between the two devices when examining the patients at the arms. In addition, in participants of 50 years and older the Vibrameter showed a considerable variability between the three obtained vibration values. Variability between the values obtained with the tuning fork was rarely seen.

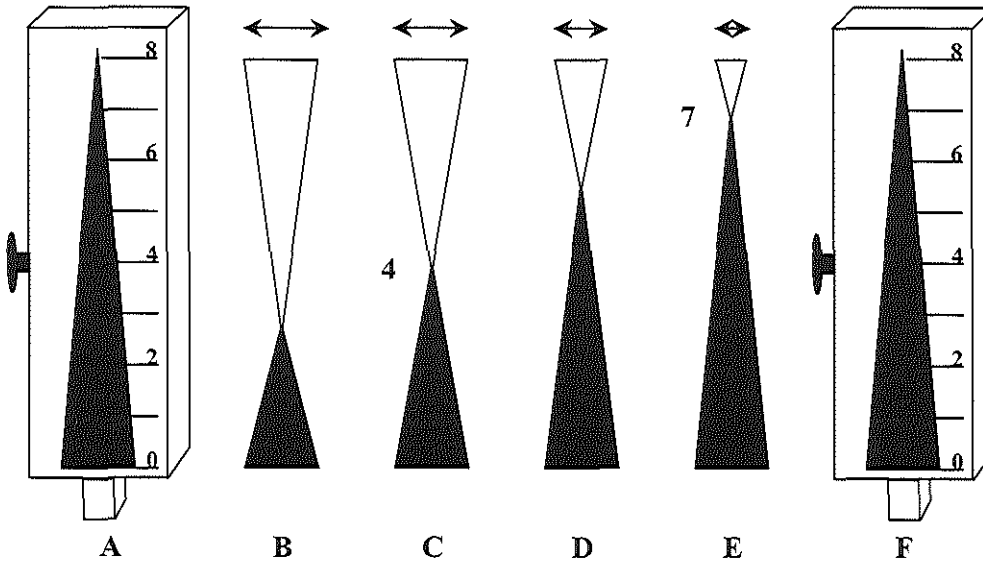
The vibration values at the four sites of examination in the healthy controls and polyneuropathy patients were combined in each participant in order to calculate new sensitivities and specificities at a chosen definition for having polyneuropathy. The 5% lower limit values obtained for the Rydel-Seiffer tuning fork and the 95th percentile vibration values for the Vibrameter were used for this purpose. Describing polyneuropathy as an abnormal vibration sense at at least one of the four sites examined with the Rydel-Seiffer tuning fork resulted in a sensitivity of 83% and a specificity of 80%. The sensitivity and specificity for the Vibrameter were 78% and 86%, respectively. The corresponding sensitivity and specificity when at least 2 sites investigated demonstrated an abnormal vibration sense was 63% and 94% for the tuning fork, and 61% and 98% for the Vibrameter.

Discussion

The present study provides clinical useful normal values for the vibration sense using the graduated Rydel-Seiffer tuning fork at four sites of examination. These normal values are different from those provided by the literature. Claus, *et al.* stated that individuals up to 40 years should score at least 6/8 and those above 40 years at least 4/8 (7). The difference may be explained by the difference in study design: their population consisted of only 73 healthy controls and no information was given regarding stratification for age and sex. Moreover, examination was performed only at the internal malleolus and thus no values were provided for the arms or the hallux. Three other studies provided normal values for only one site of examination, but unfortunately these values were presented in a graphical way and not translated for general clinical use (table 1, page 45; 8-10). Thivolet *et al.*, however, examined 88 healthy controls at 4 different sites, but only the hallux was graphically presented (8).

We found an age-related decrease in vibration sense for all four sites of examination. Pearson (16) was the first to report an age-related decrease and subsequently, numerous reports have confirmed this finding (3,5,8-10,15,17-23). The significance of this decline is not clear and neither is its cause. However, it is known that degenerative transformations of the Pacinian corpuscles, demyelinisation and fibre loss in peripheral nerves occur with

Figure 1



Legend to figure 1
Extremities of the Rydel-Seiffer graduated tuning fork at rest (A). Once the extremities are swinging, the fork vibrates at 64 Hz and the triangles on the weights appear double (B→E). The intersection of these two virtual triangles moves from 0 to 8 in an exponential way with decreasing vibration-amplitude of the arms (↔).

ageing (24). Degenerative changes also occur in the central nervous system with advancing age that may also account for the decreasing vibration sense (18,24-26). The decrease of vibration sense with age is, however, gradual and there is no trend towards absent vibration sense in healthy old persons, not even at the hallux (figure 4). We are inclined to conclude that absence of vibration sense, even in elderly people, should therefore be considered as abnormal. Many experienced investigators and neurologists make explicit references to a "non-specific neuropathy of ageing" and suggest that these changes occur uniquely during the ageing of peripheral nerves (24,26). However, interpretation of an absent distal vibration sense should be carefully taken, and preferably viewed in the context of other symptoms or signs compatible with a polyneuropathy. A thorough evaluation to identify possible causes should always be performed prior to attributing these abnormalities to ageing.

The obtained normal vibration values were higher for the arms than for the legs. These findings have also been reported by others (5,17,18). A possible explanation for these variances can be found in the differences in length of nerves between the arms and legs. It is known that longer axons are more prone to degeneration of distal regions, possibly due to a metabolic abnormality leading to failure of axonal transport and subsequent degeneration (27).

The sensitivity of the Rydel-Seiffer tuning fork for patients with a polyneuropathy is highly dependent on the selection of patients. In patients with predominantly motor signs and those with only small size myelinated/unmyelinated nerve disturbances it might be expected that

Table 1

Reported studies in which the graduated Rydel-Seiffer tuning fork was used

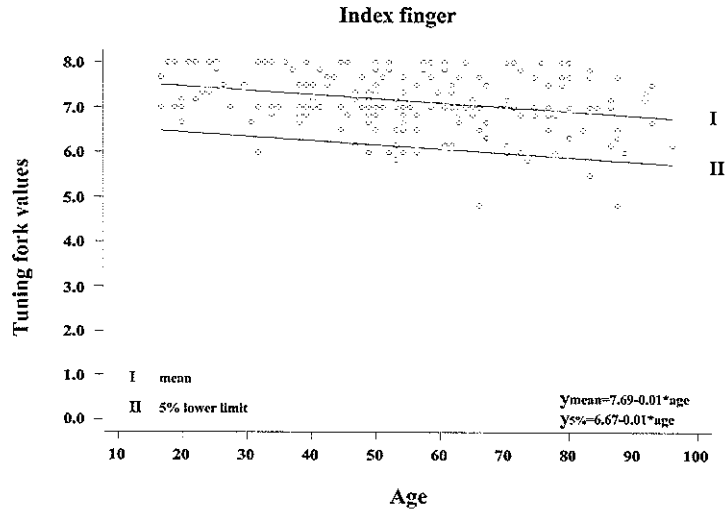
Authors (ref.)	Years	Population	site of examination	reference values	Equation
Claus *et al.* (7)	1988	73 healthy persons 26 DM pnp	internal malleolus	<40 yrs → > 6/8 >40 yrs → > 4/8	Log(VT)=0.015*age - 0.94
Crausaz *et al.* (10)	1988	177 healthy persons 89 DM patients	external malleolus	- (graphic)	-
Thivolet *et al.* (8)	1990	88 healthy persons 189 DM patients	hallux/internal malleolus/tibial crest/thumb	- (graphic)	-
Liniger *et al.* (9)	1990	214 healthy persons 192 DM patients	first metatarsal	- (graphic)	log(8.5 - VT)=0.009*age - 0.1663
Hotta *et al.* (14)	1994	57 DM patients	first three fingers	-	-
Hilz *et al.* (13)	1995	40 uremic pnp 35 alcohol pnp	internal malleolus	-	-
Hilz *et al.* (12)	1995	20 uremic pnp	internal malleolus	-	-
Bergin *et al.* (11)	1995	32 healthy persons 25 pnp (?)	internal malleolus/ tibial tuberositas	-	-

the sensitivity for detecting vibration sense abnormalities using this tuning fork will be low. Most patients with a polyneuropathy in the present study had limited sensory disturbances and hardly ambulatory problems. These findings might explain the low sensitivity that was found at the sites examined in the arms and the somewhat higher sensitivity for the two locations investigated in the legs. The vibration threshold values obtained with the Rydel-Seiffer tuning fork showed a significant correlation with those of the Vibrameter. This correlation was, however, moderate and was probably due to an increasing variability between the values obtained with the Vibrameter in participants 50 years and older. A considerable variability when applying the Vibrameter in elderly people has also been reported earlier (15,23,28). Others have shown a significant correlation between these two instruments (9-11). Hotta *et al.* found a good correlation between a variant of the graduated Rydel-Seiffer tuning fork and the Vibrameter and stated that the tuning fork provided objective and reliable measurements of vibration sense (14).

The pocket-sized Rydel-Seiffer tuning fork detected slightly more abnormalities in the selected group of patients with a polyneuropathy than the Vibrameter. The present study also demonstrates that the Rydel-Seiffer tuning fork is easily applicable, does not require long periods of attention, and measures vibration sense quicker than the Vibrameter.

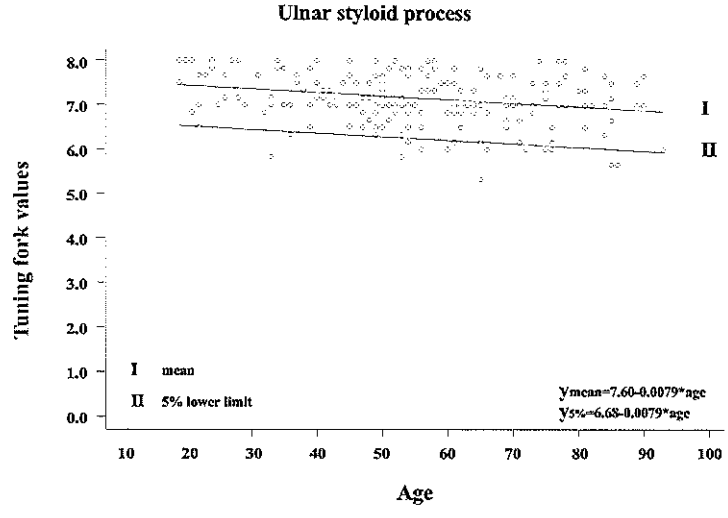
The clinical significance of a screening test depends not only on its simplicity and validity, but also on its reliability and sensitivity to register clinically relevant information. Long-term follow-up of patients with a polyneuropathy will determine the sensitivity of the Rydel-Seiffer tuning fork to changes in time and will indicate whether clinical improvement correlates with improvement in vibration sense. This is currently being evaluated more extensively in a group of patients with immune-mediated polyneuropathies with various degrees of severity. Its interobserver and intraobserver reliability and sensitivity to register changes in time are under investigation. In conclusion, our findings indicate that the Rydel-Seiffer tuning fork is a simple instrument that rapidly measures vibration sense in a semi-quantitative way. Clinical useful vibration threshold normal values are provided for this tuning fork. We propose to incorporate this pocket-sized instrument in the routine neurological examination.

Figure 2



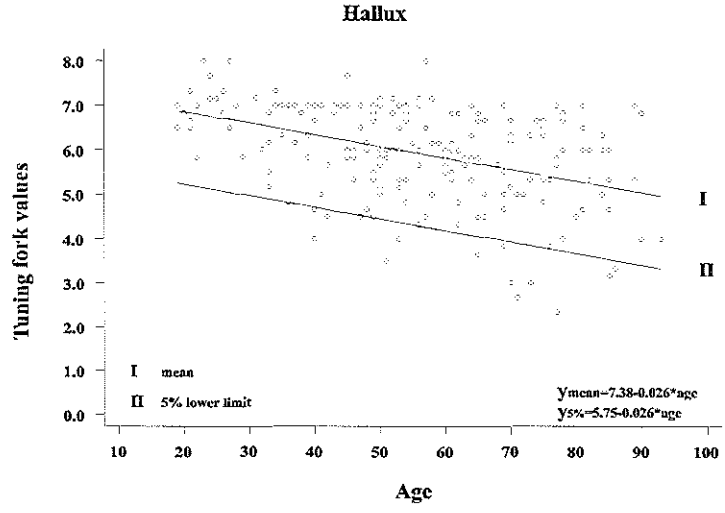
Legend to figure 2
Vibration threshold values obtained by the Rydel-Seiffer tuning fork at the Index finger. Residual standard deviation (SDres = residual standard deviation around the mean regression line) = 0.62.

Figure 3



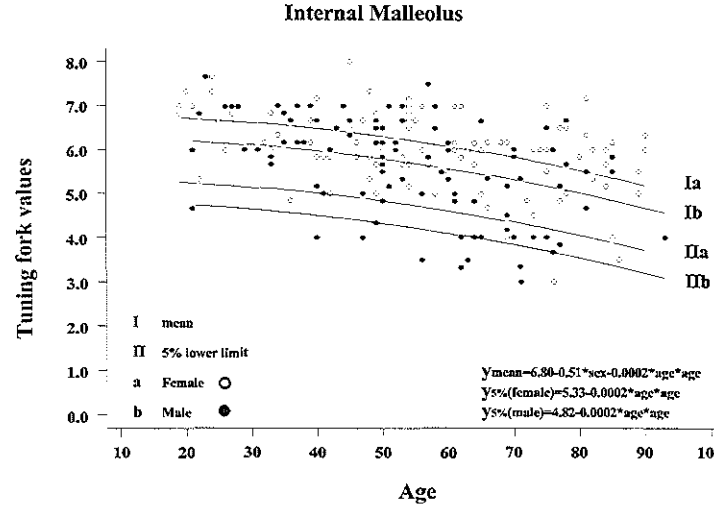
Legend to figure 3
Vibration threshold values obtained by the Rydel-Seiffer tuning fork at the Ulnar styloid process. Residual standard deviation (SDres = residual standard deviation around the mean regression line) = 0.55.

Figure 4



Legend to figure 4
Vibration threshold values obtained by the Rydel-Seiffer tuning fork at the Hallux. Residual standard deviation (SDres = residual standard deviation around the mean regression line) = 0.99.

Figure 5



Legend to figure 5
Vibration threshold values obtained by the Rydel-Seiffer tuning fork at the Internal malleolus. Residual standard deviation (SDres = residual standard deviation around the mean regression line) = 0.89. Sex = 1 (male); Sex = 0 (female).

Table 2

Normal vibration threshold values (5% lower limit) in healthy controls using the Rydel-Seiffer tuning fork

For the **upper extremities**		For the **lower extremities**	
Age (years)	Values	Age (years)	Values
≤ 40	≥ 6.5	≤ 40	≥ 4.5
41 – 85	≥ 6.0	41 - 60	≥ 4.0
> 85	≥ 5.5	61 - 85	≥ 3.5
		> 85	≥ 3.0

Legend to Table 2
Sites of examination: Dorsum of the distal interphalangeal joint of the index finger, the ulnar styloid process, dorsum of the interphalangeal joint of the hallux and the internal malleolus. Values: 0-8. The vibration threshold values are presented in rounded numbers.

Table 3

Sensitivity of the Rydel-Seiffer (RS) tuning fork compared to the Vibrameter in patients with a mild polyneuropathy (chosen specificity of 95%) (n=59)

Location	**RS tuning fork** number of patients with abnormal vibration value (%)	**Vibrameter** number of patients with abnormal vibration value (%)
Ulnar styloid process	17 (29%)	18 (31%)
Index finger	19 (32%)	19 (32%)
Internal malleolus	34 (58%)	28 (48%)
Hallux	45 (76%)	43 (73%)
Abnormal vibration sense at:		
0 site of examination	10 (17%)	13 (22%)
1 site of examination	12 (20%)	10 (17%)
2 sites of examination	19 (32%)	18 (31%)
3 or 4 sites of examination	18 (31%)	18 (31%)

References

1. Schaumburg HH, Berger AR, Thomas PK. Anatomical classification of peripheral nervous system disorders. In: Disorders of peripheral nerves. 2nd edition, Philadelphia: FA Davis Company, 1992;10-24.

2. Bickerton RC, Barr GS. The origin of the tuning fork. J R Soc Med 1987;80:771-773.

3. Bloom S, Till S, Sönksen P, *et al.* Use of a biothesiometer to measure individual vibration thresholds and their variation in 519 non-diabetic subjects. Br Med J 1984;288:1793-1795.

4. Arezzo TC, Schaumburg HH. The use of the Optacon as a screenig device. J Occup Med 1980;22:461-464.

5. Dyck PJ, Karnes J, O'Brien PC, *et al.* Detection thresholds of cutaneous sense in humans. In: Dyck PJ, Thomas PK, Griffin JW, *et al.* eds. Peripheral neuropathy. 3rd edition, Philadelphia: WB Saunders Company, 1993: 706-728.

6. Rydel A, Seiffer W. Untersuchungen über das vibrationsgefühl oder die sog. "Knochensensibilität" (Pallästhesie). Archiv für Psychiatrie und Nervenkrankenheiten 1903;37:488-536.

7. Claus D, Carvalho VP, Neundörfer B, *et al.* Zur unterzuchung des vibrationsempfindens. Nervenarzt 1988;59:138-142.

8. Thivolet C, El Farkh J, Petiot A, *et al.* Measuring vibration sense with graduated tuning fork. Diab Care 1990;13::1077-1080.

9. Liniger C, Albeanu A, Bloise D, *et al.* The tuning fork revisited. Diab Med 1990;7:859-864.

10. Crausaz FM, Clavel S, Liniger C, *et al.* Additional factors associated with plantar ulcers in diabetic neuropathy. Diab Med 1988;5:771-775.

11. Bergin PS, Bronstein AM, Murray NMF, *et al.* Body sway and vibration perception thresholds in normal ageing and in patients with polyneuropathy. J Neurol Neurosurg Psychiatry 1995;58:335-340.

12. Hilz MJ, Claus D, Rösl G, *et al.* Besserung des temperatur- und vibrationsempfindens bei chronischer urämie nach einzeldialyse. Fortschr Neurol Psychiat 1995;63:264-269.

13. Hilz MJ, Zimmerman P, Rösl G, *et al.* Vibrameter testing facilitates the diagnosis of uremic and alcoholic polyneuropathy. Acta Neurol Scand 1995;92:486-490.

14. Hotta N, Sugimura K, Tsuchida I, *et al.* Use of the C64 quantitative tuning fork and the effect of niceritrol in diabetic neuropathy. Clin Ther 1994;16:1007-1015.

15. Goldberg JM, Lindblom U. Standardised method of determining vibratory perception thresholds for diagnosing and screening in neurological investigation. J Neurol Neurosurg Psychiatry 1979;42:793-803.

16. Pearson GHJ. Effect of age on vibratory sensibility. Arch Neurol Psychiatry 1928;20:482-496.

17. De Michele G, Filla A, Coppola N, *et al.* Influence of age, gender, height and education on vibration sense. A study by tuning fork in 192 normal subjects. J Neurol Sci 1991;105:155-158.

18. Wiles PG, Pearce SM, Rice PJS, *et al.* Vibration perception threshold: Influence of age, height, sex, and smoking, and calculation of accurate centile values. Diab Med 1991;8:157-161.

19. Merchut MP, Toleikis SC. Ageing and quantitative sensory thresholds. Electromyogr Clin Neurophysiol 1990;30:293-297.

20. Critchley M. The neurology of old age. Lancet 1931;1:1221-1230.

21. Kokmen E, Bossemeyer Jr RW, Barney J, *et al.* Neurological manifestations of ageing. J Gerontology 1977;32:411-419.

22. Steinberg FU, Graber AL. The effect of age and peripheral circulation on the perception of vibration. Arch Phys Med Rehabil 1963;44:645-650.

23. Thomson FJ, Masson EA, Boulton AJM. The clinical diagnosis of sensory neuropathy in elderly people. Diab Med 1993;10:843-846.

24. Knox CA. Neuroanatomical changes associated with ageing in the peripheral nervous system. In: Albert ML, Knoefel JE eds. Clinical neurology of ageing. 2nd edition, New York: Oxford University Press, 1994: 68-78.

25. Howell TH. Senile deterioration of the central nervous system: A clinical study. Br Med J 1949;1:56-58.

26. Thomas PK. Late life neuropathy: an overview. New Iss Neurosci 1988;1:85-94.

27. Griffin JW, Watson DF. Axonal transport in neurological disease. Ann Neurol 1988;23:3-13.

28. Fagius J, Wahren LK. Variability of sensory threshold determination in clinical use. J Neurol Sci 1981;51:11-27.

Chapter 3

Reliability and responsiveness of the graduated Rydel-Seiffer tuning fork in immune-mediated polyneuropathies

Merkies ISJ[1], Schmitz PIM[2], van der Meché FGA[1], van Doorn PA[1], for the Inflammatory Neuropathy Cause And Treatment (INCAT) group[3]

From the Department of [1]Neurology, [2]Department of Statistics - Daniel den Hoed Cancer Centre, Erasmus University Medical Centre Rotterdam, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands. [3]See Chapter 2 for INCAT-members.

J Neurol Neurosurg Psychiatry, 2000;68:669-671; published with permission from the BMJ Publishing Group

Abstract

Objectives: To examine the reliability and responsiveness of the Rydel-Seiffer graduated tuning fork in immune-mediated polyneuropathies.
Methods: This tuning fork was applied in 113 patients with a clinically stable condition (83 who had had Guillain-Barré syndrome (GBS) in the past, 22 with a chronic inflammatory demyelinating polyneuropathy (CIDP), and 8 patients with a monoclonal gammopathy of undetermined significance associated polyneuropathy) and serially in 20 patients with recently diagnosed GBS (n=7) or CIDP (n=13) with changing clinical conditions. The measures were done in triplicate at eight different locations in the limbs and the values were compared with the recently published vibration threshold reference values.
Results: Good interobserver and intraobserver agreements (quadratic weighted kappa = 0.67 - 0.98) and high responsiveness values (standardised response mean scores > 0.8) were demonstrated for the Rydel-Seiffer tuning fork.
Conclusion: These results provide, in addition to literature findings, further evidence for incorporation of this easily applicable instrument in routine neurological examination.

Introduction

Clinically useful vibration threshold reference values were recently published for the Rydel-Seiffer (Martin, Tuttlingen, Germany) graduated tuning fork and its construct convergent validity was also demonstrated after correlation with an electronic device, the Vibrameter (see appendix I, page 207; reference 1). In addition, we investigated the interobserver and intraobserver reliability and responsiveness to clinical changes over time of this tuning fork in patients with immune-mediated polyneuropathies (2). Patients with Guillain-Barré syndrome (GBS), chronic inflammatory demyelinating polyneuropathy (CIDP), or a polyneuropathy associated with a monoclonal gammopathy of undetermined significance (MGUSP) were recruited, as it was argued that these disorders represent parts of a continuum regarding their neuromuscular dysfunction pattern (3).

Patients and methods

Patients

One hundred and thirteen patients (83 with GBS, 22 with CIDP, and 8 with MGUSP) with a stable neurological condition were recruited from the Rotterdam immune-mediated polyneuropathy databank and the Dutch GBS study group (stable group). A stable neurological condition was required in order to obtain the highest reliability when assessing vibration sense with the Rydel-Seiffer tuning fork. These selected patients still had residual symptoms and/or signs due to their illness, representing a broad range of disability. Nine CIDP patients required interval treatment ranging from weeks to months, with intravenous immunoglobulin (IVIg). With this therapy their clinical condition has been stable for more than 6 months.

Twenty other patients recently diagnosed with GBS (n = 7) or CIDP (n = 13) were enrolled to investigate the responsiveness of the Rydel-Seiffer tuning fork (longitudinal group). All GBS and CIDP patients met the research criteria for their illness (4,5). The diagnosis MGUSP was established after excluding all possible causes for the gammopathy and polyneuropathy (6).

Assessment tool/scale

The *Rydel-Seiffer* tuning fork (Martin, Tuttlingen, Germany) is a graduated fork that determines the ability of individuals to discriminate between various vibration intensities (see appendix I, page 207; references 1,7). A triangle and an arbitrary scale from 0 (minimum score) to 8 (maximum score) imprinted on the weights allow quantitative vibration assessment. Once the arms are swinging, the fork vibrates at 64 Hz and the triangles on the weights appear double. The intersection of these two virtual triangles moves from 0 to 8 in an exponential way with decreasing vibration-amplitude of the arms (1).

The *Overall disability sumscore* is composed by recently published arm and leg disability scales that were slightly modified, with a total score ranging from 0 ("no signs of disability") to 12 ("most severe disability score") (8). This scale was assessed to investigate the possible impact of vibration sense abnormalities on total disability. The overall disability sumscore comprises a good functional description of the arms and legs in a checklist form suitable for interviewing patients. Daily arm activities like dressing upper part of the body, doing and

undoing buttons and zips, washing and brushing hair, using a knife and fork, and turning a key in a lock are scored as being "not affected", "affected but not prevented" or "prevented". Subsequently, these results are translated into an arm grade (score range: 0 [normal arm abilities] to 5 [severe symptoms and signs in both arms preventing all purposeful movements]). The leg scale highlights problems regarding walking taking into account the use of a device. The obtained results are also translated into a leg grade (score range: 0 [walking is not affected] to 7 [restricted to wheelchair or bed most of the day, preventing all purposeful movements of the legs]) (8).

Test procedures

General aspects. All participants gave informed consent before the study. All measures were obtained in a quiet and temperature controlled room at our outpatient clinic. The patients were examined in supine position. Vibration was assessed from distal to proximal and only the most affected side and the highest extension of dysfunction were registered for the arms and legs separately. The tuning fork was applied as perpendicular as possible resting on its own weight with the arms of the fork swinging maximally. The vibration extinction threshold was considered as the nearest value (recorded as a multiple of 0.5 points) to the apparent point of intersection of the virtual triangles when the patient indicates that vibration was no longer perceived. This threshold was calculated by averaging the readings of three repeated tests. The averaged values were compared with the recently reported vibration threshold normal values and graded as follows: normal (grade = 0) or disturbed (grade = 1) vibration sense at the dorsum distal interphalangeal joint of the index finger or hallux; abnormal sense at the ulnar styloid process or medial malleolus (grade = 2), at the medial humerus epicondyle or patella (grade = 3), at acromio-clavicular joint or anterior superior iliac spine (grade = 4) (1).

Reliability. For the reliability assessment of the tuning fork, two senior neurologists and six residents in neurology formed 28 different pairs of examiners. Preceding the study, all investigators received instructions in assessing the outcome measures. Twenty-seven ("variable") couples examined a total of 68 patients (2 to 3 patients a couple). The remaining 45 patients were investigated by the "experienced" couple (couple number 1). This couple was formed to investigate the effect of training and thus a possible increase in interobserver and intraobserver reliability when using the Rydel-Seiffer tuning fork more often.

The stable patients were examined at two different occasions at our outpatient clinic. During the first visit the members of an appointed pair performed their scores independently and consecutively (usually within 2 hours) (inter-observer measures). Within 2-4 weeks, the patient returned for a second visit and only one investigator of the earlier assigned pair of researchers examined the patient again without having access to previous results (intra-observer measures). The assessments sequence at entry and the examination at the second visit were equally distributed among the members of an assigned couple. Overall, each member of a couple examined approximately the same amount of patients as their partner. The Rydel-Seiffer tuning fork and the overall disability sumscore were assessed at each visit. However, only the values of the overall disability sumscore at entry were used to determine the range of disability in this group of patients.

Responsiveness. Vibration sense and the overall disability sumscore were assessed in the longitudinal group of patients by the same clinician (ISJM) at study entry and 8-13 times in each patient during follow-up. There was a standard follow-up schedule (weeks 0, 2, 4, 8,

12, 16, 21, 26, 40, 52) with additional investigations if necessary. The study took place between March 1997 and July 1999 and was part of a comprehensive research on outcome in patients with immune-mediated polyneuropathies on behalf of the *INCAT*-group.

Statistics

The interobserver and intraobserver reliability for the obtained vibration sense grades was determined using the weighted kappa-statistic (κ) measures for the two investigator ("experienced" and "variable") groups (9). The weights of the kappa were defined as $1-[(i-j)/(k-1)]^2$ (i = rows and j = columns of the ratings by two observers, k = maximum number of possible ratings).

Responsiveness was investigated by calculating the standardised response mean (SRM) score for the total vibration grades at various arbitrarily chosen occasions during follow-up (weeks 12, 26, 40, 52) (10). The SRM is equal to the mean change in score divided by the standard deviation of the change in score ($SRM=\mu i-\mu o/SD(\mu i-\mu o)$; μi = mean vibration score of the longitudinally examined group at week = i; μo = mean vibration score at week = 0) (10). According to Cohen, an SRM value between 0.5 and 0.8 is considered moderate, and 0.8 or greater as good responsiveness (11).

In the longitudinally followed patients, random effects linear regression analyses of the overall disability sumscore values on the total vibration grades (arm [range: 0 - 4] + leg vibration grades [range: 0 - 4] = total vibration grades [range: 0 - 8] in each patient) were performed, taking into account the correlation of the data caused by the longitudinal structure. The latter was achieved using the program "xtreg" in STATA 5.0 for Windows 95 which is based upon a cross-sectional time-series regression model as described by Dwyer and Feinleib (StataCorp. 1997. Stata Statistical Software: Release 5.0. College Station, TX: Stata Corporation) (12). A logarithmic transformation was applied to the variables (total vibration grades, overall disability sumscore) before the regression studies. Finally, median total vibration grades and disability scores at 12, 26, 40, and 52 weeks of follow-up were compared with the median value at entry (Wilcoxon signed-rank test). All analyses were performed using STATA 5.0 for Windows 95. A p-value ≤ 0.05 was considered statistically significant.

Results

General aspects. The basic characteristics of all patients in the current study are presented in the table 1. The stable group of patients (54 females; 59 males; median age 56, range 14 - 84 years) had a median duration of symptoms at onset of the study of 5.1 years. The median overall disability sumscore in this group was 4 at entry (range: 0 - 11). In these patients, the median value of the total vibration grades (arm + leg vibration grades) was 1 (range: 0 - 8) at all three assessments. Seven of these patients were bed bound and fourteen patients required assistance or a device to walk short distances. The remaining patients could walk independently.

Eight females and twelve males (median age 54.0, range 15 - 70 years) were examined longitudinally (table 1). In these patients, the median overall disability sumscore was at entry

Table 1

Characteristics of patients with immune-mediated polyneuropathies

Stable group (n = 113; GBS 83, CIDP 22, MGUSP 8)		
Sex, No (%)	Females	54 (48%)
	Males	59 (52%)
Median age at start of the study (range) in years		56 (14 - 84)
Median duration of symptoms till onset of study (years)		
	Overall	5.1
	GBS	5.2
	CIDP	3.9
	MGUSP	3.6
Longitudinal group (n = 20; GBS 7, CIDP 13)		
Sex, No (%)	Females	8 (40%)
	Males	12 (60%)
Median age at start of the study (range) in years		54 (15 - 70)

5 (range: 3 - 11). At study entry, four patients were bed bound, one requiring artificial ventilation, and nine patients were unable to walk independently. The initial median total vibration grade was 4 (range: 0 - 8) in these patients.

Reliability and Responsiveness of the Rydel-Seiffer tuning fork. In the stable group, good interobserver and intraobserver reliability values were demonstrated for the Rydel-Seiffer tuning fork by the "experienced" and "variable" couples (table 2).

All longitudinally examined patients experienced during follow-up sensory disturbances including abnormalities in vibration sense as assessed with the Rydel-Seiffer tuning fork. A total of 185 visits were completed during a follow-up period of 26 - 58 weeks. With the exception of one patient with GBS who only experienced mild symptoms, all patients received initial treatment with IVIg (0.4 g/kg/day for 5 consecutive days). All but one patient with CIDP showed good functional improvement on IVIg. The non-responder received a course of treatment with oral prednisone (100 mg/day), for 4 consecutive weeks. This patient also improved with this therapy and prednisone was tapered down in 5 months period to 30 mg every other day.

The patients with GBS did not show any deterioration. After initial improvement, all 12 IVIg responsive CIDP patients showed deterioration in their clinical condition with

Table 2

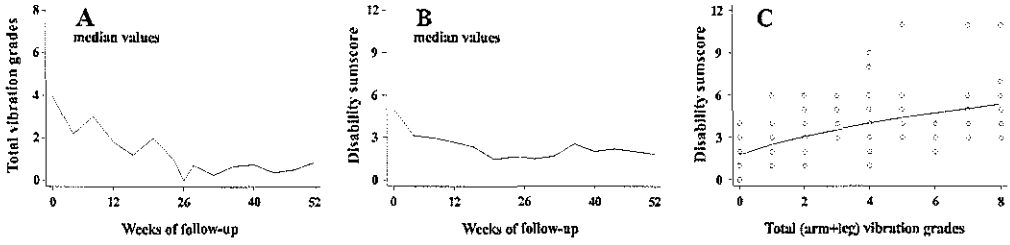
Reliability evaluation of the Rydel-Seiffer tuning fork in patients with a clinically stable immune-mediated polyneuropathy

	"Experienced" couple of examiners (couple number 1) 45 patients		"Variable" couples of examiners (couples number 2-28) 68 patients	Total (n=113)
Reliability	Weighted kappa (κ)[2]	**p-value**	Weighted kappa (κ)[2]	**p-value**
Interobserver values arms	0.67	<0.0001	0.77	<0.0001
Interobserver values legs	0.83	<0.0001	0.82	<0.0001
Intraobserver values arms	0.98	<0.0001	0.71	<0.0001
Intraobserver values legs	0.79	<0.0001	0.85	<0.0001
Interobserver values arms+legs	0.80	<0.0001	0.86	<0.0001
Intraobserver values arms+legs	0.90	<0.0001	0.84	<0.0001

increasing vibration sense abnormalities (compatible with higher vibration grades) at examination. Maintenance therapy with IVIg (1-2 days 0.4 g/kg/day at intervals of 3 - 21 weeks) was needed. Eventually, at 26 weeks of follow-up, all but one patient with GBS (19/20 = 95%) were independent in daily activities such as walking, dressing up, and eating with a knife and fork. At 1 year, the 13 patients who completed this period were also independent in daily activities with no or only minor symptoms or signs. Improvement in the longitudinal group resulted in a reduction of total vibration grades (compatible with improvement) and disability score during follow-up (figure). The median total vibration grades were lower during follow-up (2 at week 12; 0 at week 26; 0.5 at weeks 40; 1 at week 52) compared with the median value of 4.0 at entry (Wilcoxon signed-rank test: $p = 0.01$ - 0.0007). Also, as can be seen in the figure, a gradual and significant decrement in the median overall disability sumscore was also noted during follow-up (median value: 2 at week 12; 1.5 at week 26; 2 at weeks 40 and 52; compared with entry value of 5: Wilcoxon signed-rank test: $p = 0.01$ - 0.005). A significant association was demonstrated between the total vibration grades and overall disability sumscore in these patients (figure; random effects linear regression analyses; $R = 0.65$, $p < 0.0001$). Good standardised response mean scores were calculated for the RS tuning fork in these patients (SRM values: 0.9, 1.2, 1.0, and 1.2 at the weeks 12, 26, 40, and 52, respectively).

Figure

Change of total vibration grades (A) and disability sumscore (B), and their association (C)



Legend to figure
Twenty patients (7 GBS, 13 CIDP) were longitudinally examined. A total of 185 visits were completed. Total vibration grades = arm + leg vibration values and range from 0 ("no vibration abnormalities") to 8 ("most severe vibration abnormalities"). The overall disability sumscore ranges from 0 ("no signs of disability") to 12 ("most severe disability score"). The association between these two variables is expressed in (C) using regression analyses, taking into account the clustering of data at the individual patient's level ($R = 0.65$; $p < 0.0001$; see also section statistics: "xtreg") (12).

Discussion

In the current study, good interobserver and intraobserver reliability and high responsiveness are demonstrated for the Rydel-Seiffer graduated tuning fork in patients with immune-

mediated polyneuropathies. Hence, all clinimetric requirements are accomplished for this tuning fork by combining these results with literature findings (1,2,13-15). Others have also demonstrated acceptable reliability values for this instrument (13-15). However, by contrast with the current study, these reliability values were reported only for distal examination in the limbs (13-15). Thivolet and associates reported within-test variations in the arms, but unfortunately vibration sense was only assessed at the thumbs (15). In the current study, more frequent use of the Rydel-Seiffer tuning fork did not show consistently higher reliability values, thus excluding a learning effect.

A significant association was found between the total vibration grades as assessed with the Rydel-Seiffer tuning fork and the overall disability sumscore in the longitudinally examined patients. A general decrement in disability score was often accompanied by a decrease in vibration abnormalities (figure). Apparently, this convenient pocketsize instrument monitors not only the proprioceptive function, but also provides indirect information regarding the impact of vibration abnormalities on the functional abilities of patients as measured with the overall disability sumscore. Similar findings were recently reported in a study comparing the clinical picture of patients with an axonal versus demyelinating MGUSP (16). Patients with an axonal form had lower disturbances of vibration and joint position senses that corresponded with a higher functional ability (16).

Vibration sense changes over time were adequately captured using the standardised response mean score (SRM), a statistical method postulated by Liang and associates to measure responsiveness (10). To our knowledge, this paper is the first to report on responsiveness for the Rydel-Seiffer tuning fork.

With respect to the aims in the current study, some methodological issues should be addressed. First, the obtained SRM scores for the total vibration grades assessed with the Rydel-Seiffer tuning fork only showed responsiveness within one group of patients. It is not clear whether substantial discriminative responsiveness scores will be obtained for this fork when evaluating different groups of patients, for example in a trial setting comparing a placebo versus a treatment group (17). Second, a significant association was demonstrated between the total vibration grades and overall disability sumscore. The use of the Rydel-Seiffer tuning fork may therefore be suggested as an indirect indicator of general recovery. However, this should be done with some caution, because we did not determine the portion of disability due to vibration sense disturbances as assessed with this fork compared with for example muscle strength changes. Future studies are required to determine which impairment quality will have the strongest impact on disability in patients with sensory-motor immune-mediated polyneuropathies.

In conclusion, good interobserver and intraobserver reliability and responsiveness values are provided for the Rydel-Seiffer graduated tuning fork in patients with immune-mediated polyneuropathies. Hence, all clinimetric requirements are fulfilled for this tuning fork by combining these results with literature findings. The incorporation of this instrument is therefore suggested in routine neurological examination, particularly for the assessment of vibration sense in patients with polyneuropathies.

References

1. Martina ISJ, van Koningsveld R, Schmitz PIM, *et al.* Measuring vibration threshold with a graduated tuning fork in normal ageing and in patients with polyneuropathy. J Neurol Neurosurg Psychiatry 1998;65:743-747.

2. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. J Neurol Neurosurg Psychiatry 1996;60:127-130.

3. Hartung H-P, van der Meché FGA, Pollard JD. Guillain-Barré syndrome, CIDP and other chronic immune-mediated neuropathies. Curr Opin Neurol 1998;11:497-513.

4. Ad Hoc subcommittee of the American Academy of Neurology AIDS task force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

5. Asbury AK and Cornblath DR. Assessment of current diagnostic criteria for Guillain-Barré Syndrome. Ann Neurol 1990;27(suppl):S21-S24.

6. Miescher GC, Steck AJ. Paraproteinaemic neuropathies. Baillieres Clin Neurol 1996;5:219-232.

7. Rydel A and Seiffer W. Untersuchungen über das vibrationsgefühl oder die sog. "Knochensensibilität" (Pallästhesie). *Archiv für Psychiatrie und Nervenkrankenheiten* 1903;37:488-536.

8. Sharrack B, Hughes RAC, Soudain S. Guy's neurological disability scale. Neurology 1996;243(suppl2):S32.

9. Cohen J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or part credit. Psychol Bull 1968;70:213-230.

10. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopaedic evaluation. Med Care 1990;28:632-642.

11. Cohen J. Statistical power analysis for the behavioural sciences. New York, Academic Press, 1977.

12. Dwyer J, Feinleib M: Introduction to statistical models for longitudinal observation, in Dwyer J, Feinleib M, Lippert P, Hoffmeister H (eds): Statistical models for longitudinal studies of health. New York, Oxford University Press, 1992: pp 3-48.

13. Crausaz FM, Clavel S, Liniger C, *et al.* Additional factors associated with plantar ulcers in diabetic neuropathy. Diab Med 1988;5:771-775.

14. Liniger C, Albeanu A, Bloise D, *et al.* The tuning fork revisited. *Diab Care* 1990;7:859-864.

15. Thivolet C, El Farkh J, Petiot A, *et al.* Measuring vibration sensation with graduated tuning fork. Diab Care 1990;13:1077-1080.

16. Gorson KC, Ropper AH. Axonal neuropathy associated with monoclonal gammopathy of undetermined significance. J Neurol Neurosurg Psychiatry 1997;63:163-168.

17. Guyatt G, Walter S, Normal G. Measuring change over time: Assessing the usefulness of evaluative instruments. J Chron Dis 1987;40:171-178.

Chapter 4

Clinimetric evaluation of a new sensory scale in immune-mediated polyneuropathies

Merkies ISJ[1], Schmitz PIM[2], van der Meché FGA[1], van Doorn PA[1], for the Inflammatory Neuropathy Cause And Treatment (INCAT) group[3]

From the Department of [1]Neurology, [2]Department of Statistics - Daniel den Hoed Cancer Centre, Erasmus University Medical Centre Rotterdam, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands. [3]See Chapter 2 for INCAT-members.

Neurology 2000;54:943-949; published with permission from the American Academy of Neurology

Abstract

Objective: To perform a clinimetric evaluation of the inflammatory neuropathy cause and treatment (*INCAT*) sensory sumscore (ISS) in sensory-motor immune-mediated polyneuropathies. This new sensory scale was evaluated to strive for uniformity in assessing sensory deficit in these disorders.
Methods: The ISS comprises vibration and pinprick sense plus a two-point discrimination value and ranges from 0 (normal sensation) to 20 (maximum sensory deficit). Before its clinical use, a panel of expert neurologists concluded that the ISS has face and content validity. The construct validity of the ISS was investigated by correlation and regression studies with additional scales (Nine-hole peg test, Ten-metre walking test, an overall disability sumscore). All scales were applied in 113 patients with a stable neurological condition (83 patients who experienced Guillain-Barré syndrome (GBS) in the past, 22 with chronic inflammatory demyelinating polyneuropathy (CIDP), 8 patients with a monoclonal gammopathy associated polyneuropathy), and 10 patients with recently diagnosed GBS or CIDP with changing clinical conditions. Reliability of the ISS was evaluated in the stable patients. Its responsiveness was investigated in the patients examined longitudinally.
Results: A moderate to good validity was obtained for the ISS (stable group: $r = 0.38$ to 0.56; $p \leq 0.006$; longitudinal group: $R = 0.60$ to 0.82; $p \leq 0.007$, except for the association with the Ten-metre walking: $p = 0.08$). Acceptable internal consistency, and interobserver and intraobserver reliability were demonstrated for the ISS ($\alpha = 0.68$ to 0.87; $R = 0.85$ to 0.89, $p < 0.0001$). Standardised response mean scores for the ISS were high (≥ 0.8), indicating good responsiveness.
Conclusions: All clinimetric requirements are provided for the inflammatory neuropathy cause and treatment sensory sumscore. The use of this scale is therefore suggested for bedside evaluation of sensory deficit in the individual patient with a sensory-motor immune-mediated polyneuropathy as well as in clinical trials.

Introduction

The increased emphasis on accuracy in clinical neurological studies has intensified the need for the use of clinimetric well-evaluated outcome measures to quantify relevant deficits. In immune-mediated polyneuropathy studies, including patients with Guillain-Barré syndrome (GBS) or chronic inflammatory demyelinating polyneuropathy (CIDP), weakness has been assessed primarily using various valid and reliable motor scales that are based on the Medical Research Council grading system (1-4). Conversely, assessment of sensory deficit has been performed less in these patients, although sensory abnormalities may contribute to disability (5). It was argued that sensory deficit tends to be less obvious clinically and more prone to subjective interpretation compared with motor deficit (6). These suggestions may therefore explain the difficulties when assessing sensory deficit and the use of various sensory scales in immune-mediated polyneuropathy studies thus far, most of them not fulfilling all clinimetric requirements like being valid, reliable, and responsive to clinical changes over time (5,7-24).
Prompted by these observations, we constructed a new sensory scale, the inflammatory neuropathy cause and treatment (*INCAT*) sensory sumscore (ISS), to strive for uniformity in assessing, during bedside examinations, various sensory qualities representing different types of nerve fibres. This scale was constructed after a systematic literature review of all sensory methods applied in sensory-motor immune-mediated polyneuropathy clinical studies from January 1988 till January 1999. The ISS was created on behalf of the *INCAT* group, a collaborating force of European neurologists with special interest in immune-mediated polyneuropathies. These neurologists formed a panel that contributed to the formation and extensive evaluation of the ISS. Moreover, the clinimetric requirements (validity, reliability, and responsiveness) for the ISS were examined in patients with GBS, CIDP or a monoclonal gammopathy of undetermined significance related polyneuropathy (MGUSP) (25). These disorders are suggested to represent parts of a continuum regarding their neuromuscular dysfunction pattern (26). The construct validity of the ISS was calculated by correlation and regression studies with the measured values of the Nine-hole peg test (a dexterity test), the Ten-metre walking test, and an overall disability sumscore (ODSS) in all patients (27-29).

Patients and methods

Patients
A total of 113 patients (83 with GBS, 22 with CIDP, 8 with MGUSP) with a stable clinical condition were recruited from the Rotterdam immune-mediated polyneuropathy databank and the Dutch GBS study group (stable group). The selected patients still had residual symptoms or signs due to their illness, representing a broad range of disability. Nine CIDP patients required interval treatment ranging from weeks to months, with intravenous immunoglobulin (IVIg). With this therapy their clinical condition has been stable for more than 6 months. Additionally, ten patients with recently diagnosed GBS (n = 4) or CIDP (n = 6) with changing clinical conditions were enrolled to investigate the responsiveness of the ISS (longitudinal group). All GBS and CIDP patients met the international criteria for their illness (30,31). The diagnosis MGUSP was established after excluding all possible causes for the gammopathy and polyneuropathy (32).

Literature review

A systematic Medline search was performed from January 1988 till January 1999, reviewing all methods (apart from traditional neurological examination) evaluating the sensory system in clinical studies including patients with sensory-motor immune-mediated polyneuropathies. We investigated whether these scales have been formally evaluated in terms of being valid, reliable, and responsive before their use as a neurological outcome measure (25). Reports published in English that included 10 patients or more were identified using the following keywords: GBS, CIDP, acquired/idiopathic (poly)(radiculo)neuropathy, polyneuritis, gammopathy, dysimmune, paraprotein(a)emia, and monoclonal gammopathy of undetermined significance polyneuropathy.

Assessment tools/scales

The conceptual framework of the "*INCAT*" sensory sumscore (ISS) was created using a judgmental approach based on literature review and consensus of an expert panel consisting of 13 senior neurologists (all *INCAT* members) with special interest in immune-mediated neurological disorders (33). The ISS ranges from 0 ("normal sensation") to 20 ("most severe sensory deficit") and is composed by the summation of the following sensation qualities: pinprick arm grade [range: 0-4] and vibration arm grade [range: 0-4] and pinprick leg grade [range: 0-4] and vibration leg grade [range: 0-4] and two-point discrimination grade [range: 0-4]. Pinprick was tested using the sharp end of an esthesiometer. Patients were asked to indicate whether they experienced the pinprick as normal or abnormal. Paresthesia, dysesthesia, and hyperesthesia were scored as abnormal. We sought a normal reference point (e.g., the face) if a patient was experiencing problems indicating whether the pinprick was normal or not. Vibration sense was tested using the validated graduated Rydel-Seiffer tuning fork and the obtained measures were compared with the reported normative threshold values (34). The sites of examination with corresponding grades were defined as follows: normal (grade 0) or disturbed (grade 1) pinprick or vibration sense at the dorsum distal interphalangeal joint of the index finger or hallux; abnormal sense at the ulnar styloid process or medial malleolus (grade 2), at the medial humerus epicondyle or patella (grade 3), at acromio-clavicular joint or anterior superior iliac spine (grade 4). Pinprick and vibration sense examination took place from distal to proximal and only the highest extension of dysfunction of the most affected arm and leg was recorded separately for both qualities. If for example, the vibration sense was scored as abnormal at the index finger at both sides, but as normal at the styloid process, a more proximal examination was not performed. This patient would have a vibration grade score of 1 for the arms. For the two-point discrimination quality, a hand-held esthesiometer was used where the exact measurable distance in millimetres could be read on the instrument. This instrument was assessed in a 'static' manner at the ventral side, distal phalanx of the index finger, and the corresponding grades were arbitrarily chosen (grade 0: ≤4mm; grade 1: 5-9mm; grade 2: 10-14mm; grade 3: 15-19mm; grade 4: ≥20mm) (see appendix I, page 208).

The *Overall disability sumscore* (ODSS) is composed by an arm and leg disability scale that were slightly modified to obtain a total score ranging from 0 ("no signs of disability") to 12 ("most severe disability score") (29). The ODSS comprises a good functional description of the arms and legs in a checklist form suitable for interviewing patients.

The *Nine-hole peg test* and the *Ten-metre walking test* were also applied to all patients to measure focal disability (6,27,28).

Test procedures

General aspects. All participants gave informed consent before the start of the study. All measures were obtained in a quiet, temperature-controlled room (approximately 20°C) at our outpatient clinic. Sensory modalities were examined in random order with the patients lying in supine position. The graduated tuning fork was used as described previously (34). Briefly, this tuning fork was applied as perpendicular as possible, resting on its own weight with the arms of the fork swinging maximally. The participants indicated the moment when they no longer perceived the decreasing vibration stimulus. The readings of three repeated tests were averaged and considered the vibration value for that site of examination. The time to complete the ISS was recorded at each assessment (in seconds).

All patients received training in assessing the Nine-hole peg test before the start of the study to exclude any training effect. This test was performed under the prescribed standard conditions, in alternating order for both hands (6,27). Patients were also requested to walk 10 meters in a straight line at their preferential speed, using whatever aid needed (6,28). Three measures were completed for each of these tests and the corresponding time was recorded at each assessment (in seconds). For each test separately, the mean time of completion was calculated by averaging the three obtained measures.

Validity and Reliability. The *INCAT* expert panel was asked at various occasions to review the ISS and comment on its general structure and components, and to indicate whether this scale measures what it is supposed to measure, thus providing face and content validity for the ISS. Construct validity of the ISS was investigated by correlation and regression studies with additional scales (Nine-hole peg test, Ten-metre walking test, and the Overall disability sumscore).

For the assessment of reliability and construct validity of the ISS in the stable group of 113 patients, 2 senior neurologists and 6 experienced residents in neurology formed 28 different couples. Preceding the study, all investigators received instructions in assessing the outcome measures. Twenty-seven ("variable") couples investigated 68 patients (2 to 3 patients per couple). The remaining 45 stable patients were investigated by the "experienced" couple. The latter couple was formed to examine the effect of training (and thus a possible increase in reliability) when using the ISS often.

The patients were examined at two different occasions at our outpatient clinic. During the first visit the two members of an appointed pair acquired their scores independently and consecutively (usually within 2 hours; interobserver measures). Within 2 to 4 weeks, the patient returned for a second visit and only one investigator of the earlier assigned pair examined the patient (intraobserver values) without having access to the previous results. The assessment sequence at entry and the examination at the second visit were distributed equally among the members of an assigned couple. Eventually, each member of a couple examined approximately the same number of patients. All scales were assessed at each visit in all patients. For the validity studies, only the recruited scales' values at one visit were used.

Responsiveness. Ten patients were longitudinally examined by the same clinician (ISJM) and all scales were assessed at study entry and 8 to 13 times in each patient during follow-up. There was a standard follow-up schedule (week 0, 2, 4, 8, 12, 16, 21, 26, 32, 40, and 52) with additional clinical investigations if necessary. During each visit, the patients were requested to judge whether their clinical condition deteriorated (grade 1), remained stable

(grade 2), or improved (grade 3) when compared with the last visit ("clinical-judgement-scores"). At study entry, the patients reflected their clinical condition against their physical status within the two weeks before the start of the study. The study took place between March 1997 and January 1999.

Statistics

Validity, internal consistency, and interobserver and intraobserver Reliability. In the stable group, correlation between the ISS and additional scales was analysed using Spearman's rank correlation coefficient. Also, random effects linear regression analyses of the ISS on additional scales were performed in the longitudinal group, taking into account the correlation of the data caused by the longitudinal structure. The latter was achieved using the program "xtreg" in STATA 5.0 for Windows 95 (Stata Statistical Software: Release 5.0, 1997; Stata Corporation, College Station, TX), which is based upon a cross-sectional time-series regression model as described by Dwyer and Feinleib (35). A logarithmic transformation was applied to the variables (ISS, Nine-hole peg test, Ten-metre walking test, ODSS) before the regression analyses.

Cronbach's alpha coefficient was estimated for the ISS in both (stable and longitudinally followed) patients groups (36). The interobserver and intraobserver reliability for the ISS values in the stable group of patients was quantified by estimation of the intraclass correlation coefficient using a one-way random effects analysis of variance (Anova) model for the two investigator ("experienced" and "variable") groups.

Responsiveness. The association between the serially obtained ISS values and the clinical-judgement-scores in these patients was estimated using "xtreg" (35). Responsiveness was also investigated by calculating the standardised response mean (SRM) scores for the ISS at various, arbitrarily chosen occasions during follow-up (weeks 2, 4, 12, 26, 52) (37). SRM is equal to the mean change in score divided by the standard deviation of the change in score ($SRM = \mu i - \mu o / SD(\mu i - \mu o)$; μi = mean ISS value of the longitudinally examined patients at week = i; μo = mean ISS value at entry [week = 0]) (37). An SRM value between 0.5 and 0.8 is considered moderate, and ≥ 0.8 as high responsiveness (37,38). Median ISS values at 12, 26, 40, and 52 weeks of follow-up were compared with the median value at entry (Wilcoxon signed-rank test). All analyses were performed using Stata 5.0 for Windows 95. A value of $p \leq 0.05$ was considered statistically significant.

Results

General aspects, and Face and Content Validity of the ISS. The ISS was evaluated by thoroughly the expert panel of 13 neurologists and was modified according to their recommendations. Eventually, all members of this panel concluded that the ISS has face and content validity. Subsequently, this scale was assessed in the selected patients. All eight examiners who investigated the patients concluded that the ISS was administered easily. A median time of 270 (range: 84-751) seconds was needed to complete the ISS.

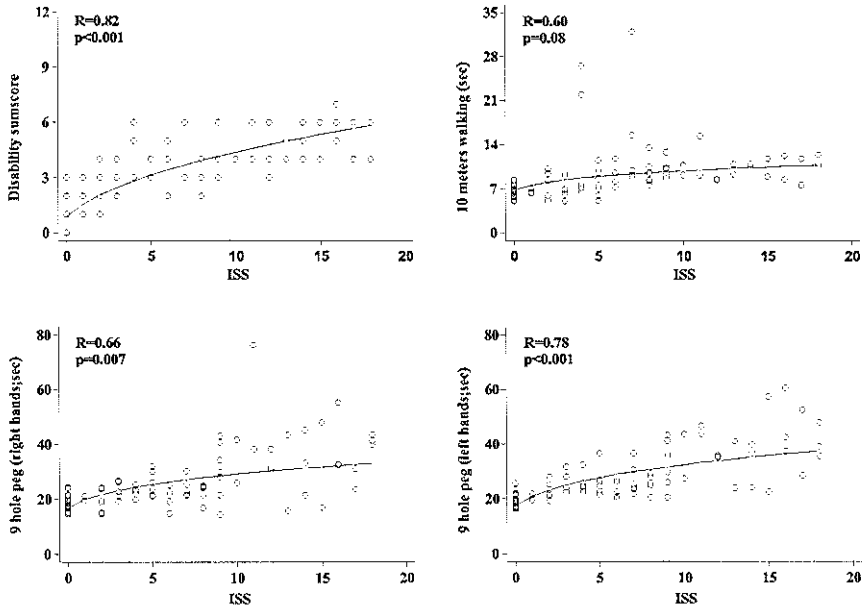
The stable group of patients (54 women; 59 men; median age 56, range 14 to 84 years) had a median duration of symptoms till onset of the study of 5.1 years. In these patients, the ISS had a median value of 3 at all three assessments (range at first assessment: 0 to 15; at second and third assessment: 0 to 18). The median time required to complete the Nine-hole peg test

in these patients was 23.1 (range: 14.6 to 143.6; for the right hands) and 23.8 seconds (range: 15.8 to 192.2; for the left hands). A median time of 8.3 (range: 5.4 to 32) seconds was needed to walk 10 meters. The overall disability sumscore ranged from 0 to 11 (median value: 4).

Construct Validity, Internal consistency, and interobserver and intraobserver Reliability of the ISS. The internal consistency value for the ISS in the stable group of patients was 0.68, 0.73, and 0.71 at first, second, and third assessments. An internal consistency value of 0.87 was obtained for the ISS in the longitudinally examined patients. The ISS was correlated moderately with the additional scales in the stable group, thus demonstrating its validity (Spearman's rank correlation coefficient; $r = 0.38$ to 0.56; $p \leq 0.006$; see Table). In the longitudinally followed patients, a significant correlation was also demonstrated between the ISS and additional scales except with the 10-meter walking test (figure 1). The interobserver and intraobserver reliability values for the ISS ranged from $R = 0.85$ to 0.89 (Anova; $p < 0.0001$ for all associations; see Table).

Figure 1

Longitudinal regressions of the "*INCAT*" sensory sumscore (ISS) on additional scales in patients with immune-mediated polyneuropathies



Legend to figure 1

Ten patients were longitudinally followed. A total of 109 visits were completed. Random effects linear regression analyses were performed of the "*INCAT*" sensory sumscore (range: 0[normal sensation] to 20[most severe sensory deficit]) on additional scales using "xtreg" (see section "Statistics"; reference 35). The Overall disability sumscore ranges from 0 ("no signs of disability") to 12 ("most severe disability").

Responsiveness of the ISS. Four women and six men (median age 40.5, range 15-70 years) were examined longitudinally. A total of 109 visits were completed in these patients. The follow-up period ranged from 52 to 58 weeks. At entry, seven patients (3 patients with GBS, 4 with CIDP) were unable to walk independently. All patients experienced general loss of strength and sensory disturbances. Except for one GBS patient, who only experienced mild symptoms, all patients received initial treatment with IVIg (0.4 g/kg/day for 5 consecutive days). During follow up, all patients showed good physical and functional improvement compared with entry. The GBS patients did not show any deterioration during follow-up. After initial improvement, all six patients with CIDP showed some clinical deterioration. Remarkably, four of these patients demonstrated at various moments an increase in sensory symptoms and signs with relatively stable muscle strength conditions (figure 2). Consequently, the intervals of IVIg therapy were shortened to regain the earlier achieved clinical and functional improvement.

Table

Reliability and validity analyses of the "*INCAT*" sensory sumscore (ISS) in a stable group of patients with immune-mediated polyneuropathies (n=113)

	"Experienced" couple of examiners (couple number 1; 45 patients)		"Variable" couples of examiners (couples number 2-28; 68 patients)	
Validity "*INCAT*" sensory sumscore versus	Spearman's rank Correlation coefficient (**r**)	**p-value**	Spearman's rank correlation coefficient (**r**)	**p-value**
Nine-hole peg test Right hands	0.47	0.002	0.42	0.0005
Nine-hole peg test Left hands	0.53	0.0003	0.49	<0.0001
Ten-metre walking time	0.55	0.0002	0.38	0.002
Overall disability sumscore	0.41	0.006	0.56	<0.0001
Reliability "*INCAT*" sensory sumscore	Intraclass correlation coefficient (**R**)	**p-value**	Intraclass correlation coefficient (**R**)	**p-value**
Interobserver agreements	0.89	<0.0001	0.86	<0.0001
Intraobserver agreements	0.85	<0.0001	0.87	<0.0001

These longitudinally followed patients graded their clinical condition 31 times as "deteriorating", 21 times as "stable", and 57 times as "improving". There was a general reduction in ISS values (compatible with improvement) during follow-up. These values were significantly associated with the clinical-judgement-scores by these patients (figure 3). The median ISS values during follow-up were lower (2, 0, 3, and 1 at the weeks 12, 26, 40, and 52) compared with the median value of 6.5 at entry (Wilcoxon signed-rank test: $p = 0.05$ to 0.006). The calculated standardised response mean (SRM) scores for the ISS were good (SRM values: 0.8, 1.0, 1.6, 2.2, and 1.1 at the weeks 2, 4, 12, 26, and 52).

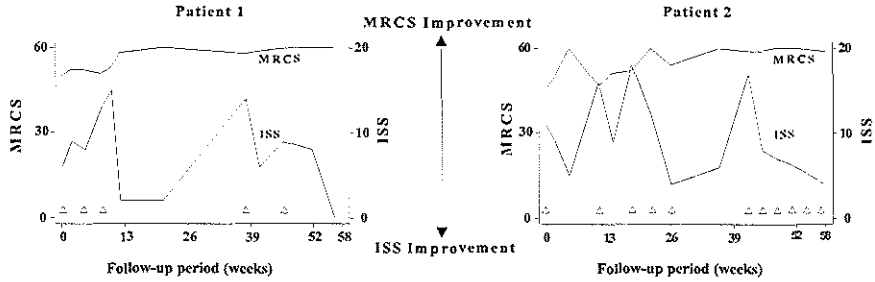
Discussion

Instruments for measuring outcome must be appropriate to the patient group being studied, must not be time-consuming, and must be easy to administer. Also, fulfilment of the clinimetric demands (being valid, reliable, and responsive to changes over time) is essential in the evaluation of such an instrument (25). In the current study, the "*INCAT*" sensory

sumscore (ISS) demonstrated to be easily applicable and only required a median time of 4.5 minutes for the evaluation of various sensory modalities in patients with sensory-motor immune-mediated polyneuropathies. Because its validity, reliability, and responsiveness were also demonstrated in these conditions, the ISS fulfils all clinimetric requirements (25). Furthermore, the reproducibility of the ISS turns out to be independent of the frequent use of

Figure 2

"*INCAT*" sensory sumscore (ISS) and MRC sumscore (MRCS) changes during follow-up in two patients with chronic inflammatory demyelinating polyneuropathy treated with intermittent intravenous immunoglobulin



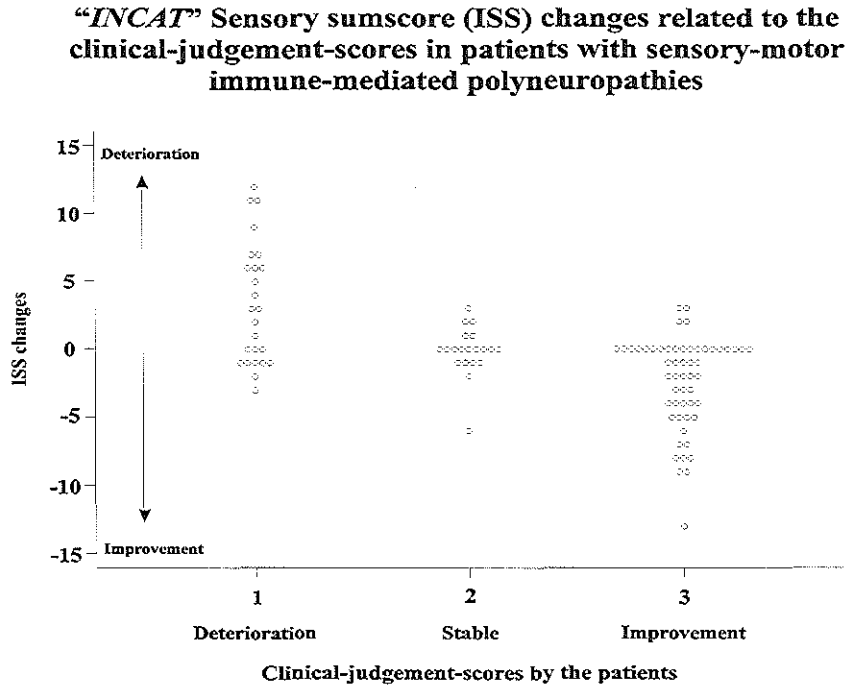
Legend to figure 2
Two patients with chronic inflammatory demyelinating polyneuropathy treated with intermittent intravenous immunoglobulin (IVIg) were selected to illustrate the prominent changes of the ISS (range: 0[normal sensation] – 20[most severe sensory deficit]) compared with muscle strength (MRC sumscore; range: 0[total paralysis] – 60[maximum strength]) (3). The triangles (Δ) indicate treatment moments with IVIg. The first treatment (week 0) consisted of 5 consecutive days of IVIg 0.4grams/kg/day. Subsequent treatments (triangles) indicate only 1 to 2 days of IVIg therapy with 0.4 grams/kg/day.

this scale, because the reliability values did not differ between the "variable" and "experienced" investigative couples.

Neurological examination has been employed traditionally to assess different sensory qualities in a wide range of disorders. During the past few decades, several attempts were made in sensory-motor immune-mediated polyneuropathy studies to quantify the results of sensory examination to reach a single score that describes a clinically detected neurological deficit. Various sensory grading systems were applied using arbitrarily defined scoring systems (5,7-24). Unfortunately, except for the sensory subset of the neurological disability score (NDS), none of these grading systems has been submitted to a comprehensive clinimetric evaluation before their general use. Regarding the NDS sensory subset, its validity and reliability were demonstrated in diabetes patients with signs of a polyneuropathy (39-41). Good internal consistency was obtained for this scale in patients with hereditary and other polyneuropathies (24). Despite these observations, the NDS has some limitations. First, the NDS does not provide information regarding the more proximal extension of sensory disturbances, because sensory qualities are only assessed at the index finger and hallux. Second, despite its general use, the NDS has not been systematically investigated in terms of its statistic and heuristic responsiveness (42). The latter is not surprising because, of

all clinimetric requirements, responsiveness has been the least studied in evaluation of outcome measures in general (42).
Clinicians and researchers need measures that discriminate between irrelevant changes (normal, minor fluctuations in the activity of an illness; "noise") and clinically meaningful changes on which a treatment policy can be based ("signal"). For this purpose, the usefulness of an outcome instrument not only depends on its simplicity, applicability,

Figure 3



Legend to figure 3
Ten patients with sensory-motor immune-mediated polyneuropathy were examined longitudinally. For each patient, the differences between all two consecutive ISS values were calculated (= ISS value at visit-*i* minus ISS value at visit-(*i-1*) = ISS changes). These differences were associated with the corresponding clinical-judgement-scores using "xtreg" (see section "Statistics"; reference 35). The circles in the figure represent the measured ISS changes for all patients. A reduction in ISS changes (compatible with less sensory disturbances over time) was associated highly with the clinical-judgement-scores by these patients (random effects linear regression analyses: $p < 0.0001$).

validity, and reliability, but also on its ability to detect these meaningful changes - an ability often addressed as "responsiveness" (42,43). A statistic and heuristic approach to examine responsiveness of a measure have been proposed (42). Statistical responsiveness captures the ability of an instrument to measure any change, irrespective of its relevance. Heuristic techniques are based on comparing changes as assessed by an outcome measure with an external indicator; for example, the clinical-judgement-scores by the patient in the current

study (42,43). We examined these two approaches in the responsiveness evaluation of the ISS. In the patients examined longitudinally, improvement of sensory disturbances (corresponding with a decrease in ISS value) was related significantly to an external criterion, the patients' clinical-judgement-scores. Additionally, the SRM scores for the ISS, a statistical measure to assess responsiveness, were equal to or higher than the proposed value of 0.8, thus representing good responsiveness (37,38,42). The ISS also helped to determine whether clinical changes over time were predominantly sensory related. As demonstrated in figure 2, substantial ISS changes were noted during follow-up whereas strength remained relatively stable. This is important, because sensory disturbances may contribute to disability (5).

With respect to the aims of the current study, some methodological issues should be addressed. First, the obtained SRM scores for the ISS only demonstrated intra-group responsiveness. It is not clear whether substantial discriminative responsiveness scores will be obtained for the ISS when evaluating various groups of patients (e.g., in a trial setting comparing a placebo versus a treated group (44). Second, a significant association was demonstrated between the ISS and disability scale. The use of the ISS may therefore be suggested as an indirect indicator of general recovery. However, this should be done with some caution, because we did not determine what portion of disability is due to sensory disturbances compared with, for example, muscle strength changes. Future studies are required to determine which impairment quality will have the strongest impact on disability in patients with sensory-motor immune-mediated polyneuropathies. Third, the arbitrarily chosen normal values for the two-point discrimination (grade = 0; $\leq$ 4 millimetres) were based primarily on the experiences of the experts panel and were not evaluated formally in terms of a possible age-related change in healthy individuals. Despite these limitations, the ISS demonstrated to be a valuable instrument for assessing sensory deficit in patients with sensory-motor immune-mediated polyneuropathies.

References

1. Medical Research Council. Aids to the investigation of the peripheral nervous system. Her Majesty's Stationery Office, London. 1943;1-2.

2. Thompson N, Choudhary P, Hughes RAC, *et al.* A novel trial design to study the effect of intravenous immunoglobulin in chronic inflammatory demyelinating polyradiculoneuropathy. J Neurol 1996;243:280-285.

3. Kleyweg RP, van der Meché FGA, Schmitz PIM. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré Syndrome. Muscle Nerve 1991;14:1103-1109.

4. Van der Meché FGA, Schmitz PIM, and the Dutch Guillain-Barré study group. A randomised trial comparing intravenous immune globulin and plasma exchange in Guillain-Barré syndrome. N Engl J Med 1992;326:1123-1129.

5. Bernsen RAJAM, Jacobs HM, de Jager AEJ, *et al.* Residual health status after Guillain-Barré syndrome. J Neurol Neurosurg Psychiatry 1997;62:637-640.

6. Wade DT. Measurements in neurological rehabilitation. Oxford: Oxford University press, 1992.

7. Gorson KC, Ropper AH, Clark BD, *et al.* Treatment of chronic inflammatory demyelinating polyneuropathy with interferon-α 2a. Neurology 1998;50:84-87.

8. Notermans NC, Wokke JHJ, Lokhorst HM, *et al.* Polyneuropathy associated with monoclonal gammopathy of undetermined significance. A prospective study of the prognostic value of clinical and laboratory abnormalities. Brain 1994;117:1385-1393.

9. Gregson NA, Koblar S, Hughes RAC. Antibodies to gangliosides in Guillain-Barré syndrome: specificity and relationship to clinical features. Q J Med 1993;86:111-117.

10. Taly AB, Veerendrakumar M, Das KB, *et al.* Sensory dysfunction in GB syndrome: A clinical and electrophysiological study of 100 patients. Electromyogr Clin Neurophysiol 1997;37:49-54.

11. Visser LH, van der Meché FGA, Meulstee J, *et al.* Cytomegalovirus infection and Guillain-Barré syndrome: the clinical, electrophysiologic, and prognostic features. Neurology 1996;47:668-673.

12. Visser LH, van der Meché FGA, Meulstee J, *et al.* Risk factors for treatment related clinical fluctuations in Guillain-Barré syndrome. J Neurol Neurosurg Psychiatry 1998;64:242-244.

13. Gorson KC, Ropper AH. Axonal neuropathy associated with monoclonal gammopathy of undetermined significance. J Neurol Neurosurg Psychiatry 1997;63:163-168.

14. Notermans NC, Lokhorst HM, Franssen H, *et al.* Intermittent cyclophosphamide and prednisone treatment of polyneuropathy associated with monoclonal gammopathy of undetermined significance. Neurology 1996;47:1227-1233.

15. Mariette X, Chastang C, Clavelou P, *et al.* A randomised clinical trial comparing interferon-α and intravenous immunoglobulin in polyneuropathy associated with monoclonal IgM. J Neurol Neurosurg Psychiatry 1997;63:28-34.

16. Hahn AF, Bolton CF, Zochodne D, *et al.* Intravenous immunoglobulin treatment in chronic inflammatory demyelinating polyneuropathy. A double-blind, placebo-controlled, cross-over study. Brain 1996;119:1067-1077.

17. Hahn AF, Bolton CF, Pillay N, *et al.* Plasma-exchange therapy in chronic inflammatory demyelinating polyneuropathy. A double-blind, sham-controlled, cross-over study. Brain 1996;119:1055-1066.

18. Dyck PJ, Low PA, Windebank AJ, *et al.* Plasma exchange in polyneuropathy associated with monoclonal gammopathy of undetermined significance. N Engl J Med 1991:325:1482-1486.

19. Suarez GA, Kelly Jr. JJ. Polyneuropathy associated with monoclonal gammopathy of undetermined significance: Further evidence that IgM-MGUS neuropathies are different than IgG-MGUS. Neurology 1993;43:1304-1308.

20. Gosselin S, Kyle RA, Dyck PJ. Neuropathy associated with monoclonal gammopathies of undetermined significance. Ann Neurol 1991;30:54-61.

21. Dyck PJ, Litchy WJ, Kratz KM, *et al.* A plasma exchange versus immune globulin infusion trial in chronic inflammatory demyelinating polyradiculoneuropathy. Annals Neurol 1994;36:838-845.

22. Oksenhendler E, Chevret S, Léger J-M, *et al.* Plasma exchange and chlorambucil in polyneuropathy associated with monoclonal IgM gammopathy. J Neurol Neurosurg Psychiatry 1995;59:243-247.

23. Dalakas MC, Quarles RH, Farrer RG, *et al.* A controlled study of intravenous immunoglobulin in demyelinating neuropathy with IgM gammopathy. Ann Neurol 1996;40:792-795.

24. Molenaar DSM, Vermeulen M, de Visser M, *et al.* Impact of neurologic signs and symptoms on functional status in peripheral neuropathies. Neurology 1999;52:151-156.

25. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. J Neurol Neurosurg Psychiatry 1996;60:127-130.

26. Hartung H-P, van der Meché FGA, Pollard JD. Guillain-Barré syndrome, CIDP, and other chronic immune-mediated neuropathies. Curr Opin Neurol 1998;11:497-513.

27. Mathiowetz V, Weber K, Kashman N, *et al.* Adult norms for the nine-hole peg test of finger dexterity. Occup Ther J Res 1985;5:24-38.

28. Collen FM, Wade DT, Bradshaw CM. Mobility after stroke: reliability of measures of impairment and disability. Int Disab Studies 1990;12:6-9.

29. Sharrack B, Hughes RAC, Soudain S. Guy's neurological disability scale [abstract]. Neurology 1996;243(suppl 2):S32.

30. Asbury AK, Cornblath DR. Assessment of current diagnostic criteria for Guillain-Barré Syndrome. Ann Neurol 1990;27(suppl):S21-S24.

31. Ad Hoc subcommittee of the American Academy of Neurology AIDS task force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

32. Miescher GC, Steck AJ. Paraproteinaemic neuropathies. Baillieres Clin Neurol 1996;5:219-232.

33. Bombardier C, Tugwell P. A methodological framework to develop and select indices for clinical trials: Statistical and judgmental approaches. J Rheumatol 1982;9:753-757.

34. Martina ISJ, van Koningsveld R, Schmitz PIM, *et al.* Measuring vibration threshold with a graduated tuning fork in normal ageing and in patients with polyneuropathy. J Neurol Neurosurg Psychiatry 1998;65:743-747.

35. Dwyer J, Feinleib M. Introduction to statistical models for longitudinal observation. In: Dwyer J, Feinleib M, Lippert P, Hoffmeister H, eds. Statistical models for longitudinal studies of health. New York, Oxford University Press, 1992: 3-48.

36. Nunnally JC. Psychometric theory. New York: McGraw Hill, 1978.

37. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;28:632-642.

38. Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press, 1977.

39. Dyck PJ, Sherman WR, Hallcher LM, *et al.* Human diabetic endoneurial sorbitol, fructose, and myo-inositol related to sural nerve morphometry. Ann Neurol 1980;8:590-596.

40. Dyck PJ, Kratz KM, Lehman KA, *et al.* The Rochester diabetic neuropathy study: Design, criteria for types of neuropathy, selection bias, and reproducibility of neuropathic tests. Neurology 1991;41:799-807.

41. Dyck PJ, Karnes JL, O'Brien PC, *et al.* The Rochester diabetic neuropathy study: Reassessment of tests and criteria for diagnosis and staged severity. Neurology 1992;42:1164-1170.

42. Liang MH. Evaluating measurement responsiveness. J Rheumatol 1995;22:1191-1192.

43. Fortin PR, Stucki G, Katz JN. Measuring relevant change: an emerging challenge in rheumatologic clinical trials. Arthritis Rheum 1995;38:1027-1030.

44. Guyatt G, Walter S, Norman G. Measuring change over time: Assessing the usefulness of evaluative instruments. J Chron Dis 1987;40:171-178.

Chapter 5

Assessing grip strength in healthy individuals and patients with immune-mediated polyneuropathies

Merkies ISJ[1], Schmitz PIM[2], Samijn JPA[1], van der Meché FGA[1], Toyka KV[3], van Doorn PA[1], for the Inflammatory Neuropathy Cause And Treatment (INCAT) group[4]

From the [1]Department of Neurology, [2]Department of Statistics - Daniel den Hoed Cancer Centre, University Hospital Rotterdam/Erasmus University, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands. From the [3]Department of Neurology, Bayerische Julius-Maximilians-Universität, Josef-Schneider Strasse 11, D-97080 Würzburg, Germany. [4]See Chapter 2 for INCAT-members.

Muscle & Nerve 2000;23:1393-1401; published with permission from Muscle & Nerve (John Wiley & Sons Inc.)

Abstract

Objectives: To provide simple and clinically useful grip strength reference values using the hand-held Vigorimeter and to examine its validity, reliability, and responsiveness in patients with immune-mediated polyneuropathies.
Methods: The Vigorimeter was applied in 530 healthy controls aged 5-93 years and in 113 patients with a stable neurological condition (83 who experienced Guillain-Barré syndrome [GBS] in the past, 22 with chronic inflammatory demyelinating polyneuropathy [CIDP], 8 with a polyneuropathy associated with a gammopathy of undetermined significance) (stable group). Additionally, this instrument was utilised serially in twenty patients with recently diagnosed GBS or CIDP and changing clinical conditions (longitudinal group). An arm disability scale was also assessed in all patients.
Results: Graphical grip strength reference values were calculated depending primarily on age and gender. Significant association was obtained between the Vigorimeter and the arm disability scale values, thus demonstrating the validity of the Vigorimeter (in stable group, Spearman rank test: $r=-0.52$ to -0.62; $p\leq0.0005$; in longitudinal group, random effect linear regression analyses: $R=0.62$-0.64, $p<0.0001$). In the stable group, good inter-/intra-observer agreements were demonstrated for the Vigorimeter (Anova: $R=0.95$-0.97). Standardised response mean (SRM) scores were high in the longitudinal group, indicating good responsiveness for this device ($SRM\geq0.8$).
Conclusions: This study provides clinically useful grip strength reference values using the Vigorimeter. The validity, reliability, and responsiveness are provided for this easily applicable instrument in patients with immune-mediated polyneuropathies. The results emphasise the value of the Vigorimeter in assessing outcome at the impairment level and indirectly at the disability level due to its association with the arm disability scale.

Introduction

The Medical Research Council (MRC) scale is primarily being used to assess muscle strength in neuromuscular disease (1). Interpreting assessments by the MRC scale is sometimes complex, because this scale has a non-linear pattern with a broad range. Perhaps, the biggest disadvantage of the MRC scale is that not all muscles can be appropriately measured using this instrument. An example is assessing strength of the small muscles embracing hand function. This is particularly of importance when distal weakness predominates, as generally occurs in patients with polyneuropathies like Guillain-Barré syndrome (GBS) and chronic inflammatory demyelinating polyneuropathy (CIDP). Especially, when evaluating therapeutic responses at the impairment level, assessment of muscles involving hand function gives important information. Grip strength, reflecting distal strength and upper limb function, is a prognostic indicator of clinical and functional recovery and is useful in monitoring the effect of treatment (2-6).

Among the instruments developed to measure grip strength quantitatively, portable dynamometers are the most popular among neurologists and rehabilitation physicians. One of these instruments is the Vigorimeter (Martin, Tuttlingen, Germany) (7,8). Reference values for this instrument have been provided for right-handed healthy individuals (9-15). However, in these studies various age groups were investigated with varying methodologies. For example, anthropometric data were not systematically examined, number of trials for each hand and resting periods between trials varied considerably, and various arm/hand/body positions were used. These observations hamper the general applicability of the obtained normal values. To date, no study has covered the evaluation of grip strength by the Vigorimeter in ages ranging from childhood to the elderly, and only a limited number of papers have addressed the clinical usefulness of this instrument (2,3,6,16,17). The Vigorimeter has been applied only twice in patients with immune-mediated polyneuropathies (2,6). A good correlation was obtained between grip strength and MRC sumscore values in a group of 11 patients with GBS, thus demonstrating the validity of the Vigorimeter (2). However, a more extensive clinical evaluation of the Vigorimeter is required in patients with immune-mediated polyneuropathies before its general use as an impairment outcome measure in these disorders is recommended (18-20).

Prompted by these considerations, we collected grip strength values and personal variables (length, weight, hand-circumference, and hand-dominance) in a large group of healthy controls with a wide age span to determine new reference grip strength values. Additionally, the validity, reliability, and responsiveness of the Vigorimeter were investigated in patients with GBS, CIDP or a polyneuropathy associated with a monoclonal gammopathy of undetermined significance (MGUSP) (18,19,21,22). The obtained grip strength values in the patients group were related to an arm disability scale (23). This was essential, not only for the validation of the Vigorimeter, but also to investigate whether grip strength could be used as an indicator of disease activity at the disability outcome level (20).

Participants and methods

Healthy controls

We recruited 530 healthy controls from hospital personnel, relatives and friends accompanying patients at our outpatient clinic, and healthy elderly, living in the village of Nieuw-Vennep as part of the municipality of Haarlemmermeer, The Netherlands. There was a door-to-door mailing and an article was published in a local newspaper (see addendum, page 86), explaining the purpose and significance of the study. Attempts were made to obtain participants representing a wide variety of social and occupational backgrounds. Children and adolescents were recruited at one primary and one high school. The eligibility criteria were clear conscious, independence in activities of daily living, socially active, sufficient vision, absence of any impairment affecting upper limb function, and normal physical development in children as reported by parents. Neurological examination was performed in all participants with special regard to the upper limb functions. Healthy individuals were stratified for age and gender and 10 age groups (5-9, 10-14, 15-19, 20-29, ..., 70-79, ≥ 80) were formed, each group consisting of 50-55 participants. The first three age groups had a smaller age-range to register the rapid change in grip strength during growth more accurately.

Patients

We recruited 113 clinically stable patients (83 with GBS, 22 with CIDP, 8 with MGUSP) from the Rotterdam immune-mediated polyneuropathy databank and the Dutch GBS study group (stable group). These selected patients still had residual signs or symptoms of their illness, representing a broad range of disability. Nine CIDP patients required interval treatment ranging from weeks to months, with intravenous immunoglobulin (IVIg). With this therapy their clinical condition was stable for more than 6 months.

Twenty consecutive patients with recently diagnosed GBS (n = 7) or CIDP (n = 13) were enrolled to investigate the responsiveness of the Vigorimeter to clinical changes over time (longitudinal group). All patients with GBS and CIDP met the international research criteria for their illness (24,25). The diagnosis MGUSP was established after excluding all possible underlying causes for the gammopathy and polyneuropathy (26). Patients were excluded from participation if there was any concomitant disease that might interfere with general strength and physical functioning.

Assessment tools/scales

The *Vigorimeter* (Martin, Tuttlingen, Germany) is an instrument that has been used to measure grip strength although, strictly speaking, it measures the air-pressure in the bulb and not force (see appendix I, page 206) (7-14, 17). The pressure in the bulb is registered on a manometer via a rubber junction tube and expressed in kiloPascals (kPa). According to the manufacturer's recommendation, the small bulb was used in the ages up to 10 years and the medium-sized bulb in the remaining participants.

The *arm disability scale* comprises a good functional description of the arms in a checklist form suitable for interviewing patients (23). Daily arm activities like dressing the upper part of the body, doing/undoing buttons and zips, washing and brushing hair, using a knife and fork, and turning a key in a lock are scored as being "not affected", "affected but not prevented" or "prevented". Subsequently, these results are translated into an arm grade. The

score range is from 0 (normal arm abilities) to 5 (severe symptoms and signs in both arms preventing all purposeful movements). Hence, this scale provides a general score for both arms. The arm disability scale is a subset of the more comprehensive Guy's neurological disability scale (23).

Test procedures

General aspects. All participants including the parents of the children gave informed consent for the study. All assessments took place in a quiet and temperature-controlled room (approximately 20° C) during daytime. Participants were examined according to the assessment recommendations by the American Society of Hand Therapists (27). All children were seated in appropriately sized chairs that allowed their feet to be flat on the floor. The investigator placed the child in the standardised position and encouraged the child to remain in that position. The bulb was positioned in the palm of the participant's hand with the air tube extending out between the individual's thumb and index finger, and with the fingers wrapped around the bulb so that the fingers touched the surface of the bulb as much as possible. Three grip strength measurements with maximum voluntary contractions for each hand were taken in alternating order. Between each trial a pause of 30 seconds was assigned. The results of three trials for each hand were averaged and considered the grip strength score for that particular hand. Half of the controls in each age group started with right hand grip assessment, and the other half with left hand grip measures, to control for the effect of order. Healthy controls were interviewed before using the Vigorimeter by one investigator (ISJM). Anthropometric data (body weight, height, hand-circumference) and hand-dominance were collected. Hand circumference was assessed using a measuring tape that was circularly applied at the distal ends of the metacarpals of digit 2 to 5. For the purpose of this study, hand-dominance was defined as the preferred hand for use. In children, hand dominance was determined by asking the participants to throw a ball or use a writing implement. Examination of the controls took place at our outpatient clinic or at the homes of the elderly; children were investigated at their schools.

Validity and Reliability. Twenty-eight different pairs of examiners were formed by two neurologists and six residents in neurology for the assessment of validity and reliability of the Vigorimeter in the stable group of 113 patients. Preceding the study, all investigators received instructions in assessing the outcome measures. Twenty-seven ("variable") couples examined 68 patients (2 to 3 patients per couple). The remaining 45 patients were investigated by the "experienced" couple (ISJM + JPAS). These two authors investigated the effect of training and thus a possible increase in interobserver and intraobserver reliability when using the Vigorimeter more often.

The patients were examined on two different occasions at our outpatient clinic. During the first visit the two members of an appointed pair performed their scores independently and consecutively (usually within 2 hours) (interobserver measures). Within 2 to 4 weeks, the patient returned for a second visit and one investigator of the earlier assigned pair examined the patient again, without having access to previous results (intraobserver values). The assessments sequence at entry and the examination at the second visit were equally distributed among members of an assigned couple. Overall, each member of a couple examined the same number of patients as their partner. The scales were assessed at each visit in all patients. For the validity studies, only the recruited scales' values at one visit were used.

Responsiveness. Twenty patients were longitudinally examined by the same clinician (ISJM) and the Vigorimeter and arm disability scale were assessed at study entry and 8 to 13 times in each patient during follow-up. There was a standard follow-up schedule (week 0, 2, 4, 8, 12, 16, 21, 26, 32, 40, and 52) with additional clinical investigations if necessary. At each visit, the patients were requested to judge whether their clinical condition deteriorated (grade 1), remained stable (grade 2) or improved (grade 3) when compared with the last visit ("clinical-judgement scores"). At study entry, the patients reflected their clinical condition against their physical status within the two weeks before the start of the study. The study took place between March 1997 and January 1999 and was part of a more comprehensive research on clinical outcome measures in patients with immune-mediated polyneuropathies on behalf of the *INCAT*-group.

Statistics

Reference values. Grip strength reference values (median and 0.05 quantile values, corresponding to a specificity of 95%) were calculated in the healthy controls, depending primarily on age and sex using quantile regression analyses with restricted cubic spline functions on age (28,29). These methods were performed to overcome the skewed distribution of grip strength values conditional on age and sex and to identify the regression curves that best fitted the data (28,29). Quantile regression analyses were also performed in the control group to determine which personal variable(s), in addition to age and gender, predict grip strength the best. These regressions were performed in each gender for the dominant and non-dominant hands separately. These grip strength values in each gender were the dependent variables and age, hand-circumference, length, and weight were the independent variables.

Validity and Reliability. In the stable group, correlation between grip strength and arm disability scale was analysed using Spearman's rank correlation coefficient. Also, random effects linear regression analyses of grip strength on the arm disability scale values was performed in the longitudinal group, taking into account the association of the data caused by the longitudinal structure. The latter was achieved using the program "xtreg" in STATA 6.0 (StataCorp. 1997. Stata Statistical Software: Release 5.0. College Station, TX), which is based upon a cross-sectional time-series regression model as described by Dwyer and Feinleib (30). The interobserver and intraobserver reliability for the obtained grip strength values in the stable group of patients was quantified by estimation of the intraclass correlation coefficient using a one-way random effects analysis-of-variance model for the two investigator ("experienced" and "variable") groups.

Responsiveness. Responsiveness was investigated by calculating the standardised response mean (SRM) scores for grip strength at various arbitrarily chosen occasions during follow-up (weeks 4, 12, 26, 40, 52) (31). SRM is equal to the mean change in score divided by the standard deviation of the change in score (SRM = μi - μo/SD(μi - μo); μi = mean grip strength value of the longitudinally examined group at week = i; μo = mean grip strength value at week = 0 [entry]) (31). A value between 0.5 and 0.8 is considered moderate, and ≥ 0.8 as high responsiveness (31,32). For each patient, the differences between all two consecutive grip strength values were calculated (grip strength at visit-i minus grip strength at visit-$(i-1)$). These differences were associated with the corresponding clinical-judgement scores (deterioration, stable or improvement) using "xtreg" (30). All analyses were

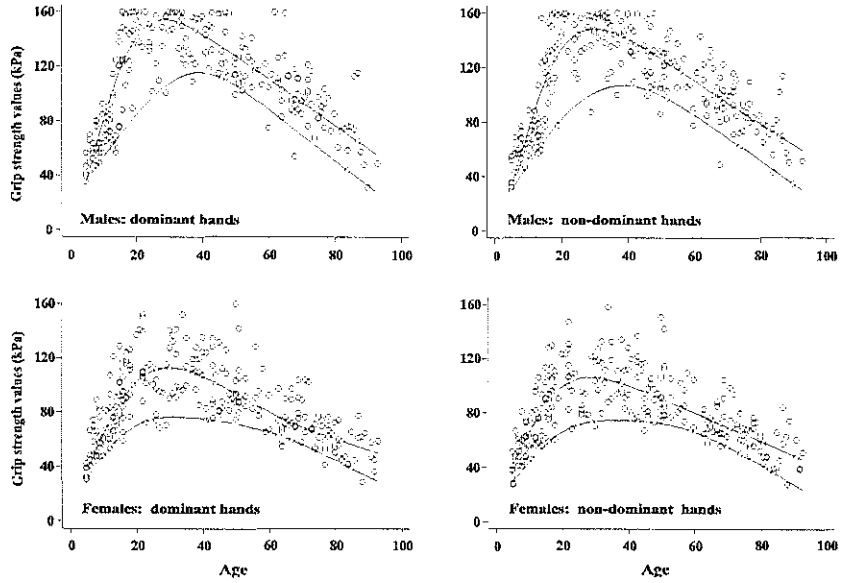
performed using Stata 6.0 for Windows 95. A value of $p \leq 0.05$ was considered statistically significant.

Results

General aspects. The whole procedure of interviewing, recruiting personal data, and assessing grip strength with the Vigorimeter took 5 to 10 minutes in each healthy participant. According to the investigators and participants, the Vigorimeter was easily administered and used.

Figure 1

Grip strength values with the Vigorimeter for the dominant and non-dominant hands in healthy males and females



Legend to Figure 1
A total of 255 healthy males (89% right-handed) and 275 health females (92% right-handed) were examined. The upper lines in each graph correspond to the calculated median grip strength values. The lower lines in each graph represent the 0.05 quantile reference values. These values were obtained in each gender separately using quantile regressions with restricted cubic spline functions (28,29).

Reference values for the Vigorimeter. Of the 551 healthy controls that were interviewed, 530 (275 females; 255 males) were enrolled in the study. Twenty-one individuals were excluded from participation, based on focal abnormalities (e.g. deformities of the forearm due to fracture, rheumatoid arthritis, Dupuytren's contracture, hemiparesis, and signs of a

polyneuropathy). The descriptive data of the healthy participants are presented in table 1. Most individuals were right-handed (91%). There were almost no differences in grip strength values between the dominant and non-dominant hands in the whole group of healthy controls (overall median differences: in healthy males 1 kPa, in healthy females 3 kPa) (figure 1). Based on these findings, only one graph was further developed for each gender separately to obtain reference values for both hands (table 2). Maximum median grip strength was observed in either gender at around 30 years. Males were consistently stronger than females (table 2). Anthropometric variables (hand-circumference, length, and weight) were consistently higher among men than women (table 1). Quantile regression analyses showed that the median grip strength values for both sides were associated significantly with age in both gender and with hand-circumference in men only ($p < 0.0001$).

Table 1

Descriptive data of the healthy controls

Variables	Males			Females		
Number of participants	n=44	n=40	n=171	n=48	n=41	n=186
Age range [years]; mean (SD)	5-12 8.5 (2.4)	13-19 15.7 (1.9)	20-93 52.6 (19.9)	5-12 8.5 (2.4)	13-19 15.6 (1.7)	20-93 54.2 (20.4)
Hand-dominance Right	40 (91%)	32 (80%)	156 (91%)	42 (88%)	35 (85%)	175 (94%)
Left	4 (9%)	8 (20%)	15 (9%)	6 (12%)	6 (15%)	11 (6%)
Length: mean (SD) range (cm)	138.0 (15.7) 108-168	178.4 (10.4) 155-195	178.9 (8.3) 159-203	137.5 (15.7) 111-169	166.8 (9.3) 131-181	165.5 (7.8) 146-186
Weight: mean (SD) range (kg)	32.0 (10.5) 17-60	62.1(11.6) 37-85	79.1 (9.8) 55-108	32.1 (11.7) 16-62	56.9 (10.0) 32-75	66.8 (10.2) 46-104
Hand-circumference: mean (SD) range (cm)	17.1 (1.6) 14-20.5	21.1 (1.4) 17.7-23.5	22.5 (1.0) 20.5-26.5	16.3 (1.6) 13.8-20.5	18.8 (1.0) 16-20.5	19.7 (1.0) 16.5-22.3

Validity and Reliability of the Vigorimeter. The stable group of patients (54 females; 59 males; median age 56, range 14-84 years) had a median duration of symptoms till study onset of 5.1 years. Ninety six percent of the patients were right-handed. The median right-hand grip strength value in this group was 67 kPa at first and third assessments (range at first assessment: 0 to 156; at third assessment: 0 to 152) and 65 (range: 0 to 158) kPa at second assessment. The median left-hand grip strength value was 65 (range: 0 to 158), 62 (range: 0 to 160), and 62 (range: 0 to 158) kPa at first, second, and third assessment. The median arm disability scale was 2 at entry (range: 0 to 4). Seven patients were bed bound and 14 required assistance or a device to walk over short distances. The remaining 92 patients could walk independently.

A moderate to good correlation was found between the Vigorimeter and the arm disability scale (ADS) values, thus demonstrating the validity of the Vigorimeter in this group of patients ("experienced" couple: grip strength right hands versus ADS: $r = -0.54$, $p = 0.0001$; left hands versus ADS: $r = -0.52$, $p = 0.0003$; "variable" couples: grip strength right hands versus ADS: $r = -0.57$, $p < 0.0001$; left hands versus ADS: $r = -0.62$, $p < 0.0001$). High interobserver and intraobserver agreements were also demonstrated for the Vigorimeter ("experienced" couple: inter-observer grip strength right hands: $R = 0.97$, left hands: $R = 0.96$; intraobserver right hands: $R = 0.96$, left hands: $R = 0.95$; "variable" couples: inter- and

Table 2

Normal values for quantitative grip strength assessment with the Vigorimeter for both hands in healthy men and women (kiloPascal)

Age (years)	Men		Women	
	0.05 quantile values per age-span	Median values per age-span	0.05 quantile values per age-span	Median values per age-span
5 – 9	42	53	38	48
10 – 14	59	95	53	73
15 – 19	72	122	63	90
20 – 24	89	146	72	106
25 – 29	101	154	76	112
30 – 34	110	153	76	112
35 – 39	115	148	76	109
40 – 44	115	141	74	104
45 – 49	111	134	73	98
50 – 54	105	126	71	92
55 – 59	94	116	67	84
60 – 64	83	106	63	76
65 – 69	74	97	58	71
70 – 74	66	91	54	67
75 – 79	57	82	48	63
80 – 84	50	75	43	59
≥ 85	37	64	35	54

intraobserver values for both hand sides: R = 0.97, except for intraobserver left hands: R = 0.96; $p < 0.0001$ for all agreements).

Responsiveness of the Vigorimeter. Eight females and twelve males (median age 54.0, range 15 to 70 years) were examined longitudinally. They all were right-handed. At study entry, four patients were bed bound, one requiring artificial ventilation, and nine patients were unable to walk independently. All patients experienced general loss of strength. A total of 201 visits were completed during a follow-up period of 40 to 58 (median: 52) weeks. Nineteen patients completed a one-year follow-up. Except for one GBS patient who only experienced mild symptoms, all patients received initial treatment with IVIg (0.4 g/kg/day for 5 consecutive days). All but one patient with CIDP showed good functional improvement on IVIg during follow-up. The non-responder received a treatment course of oral prednisone, 100 mg/day for four consecutive weeks. This patient improved also with this therapy and prednisone was tapered down over five months period to 30 mg on alternate days.

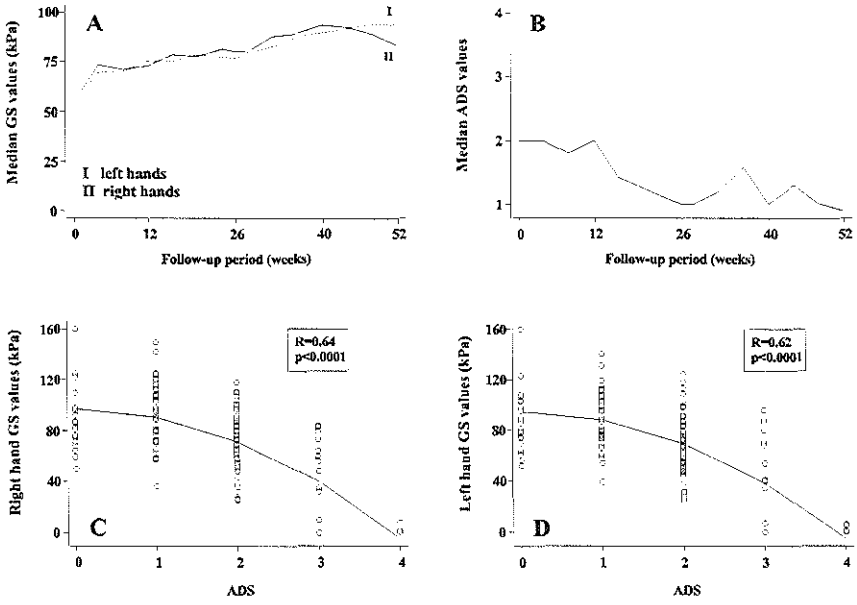
The GBS patients did not show any deterioration during follow-up. After initial improvement, all 12 IVIg responsive CIDP patients showed deterioration in their clinical condition. Consequently, maintenance therapy with IVIg (1 to 2 days 0.4 g/kg/day; intervals: 3 to 21 weeks) was needed to prevent further deterioration and to regain earlier achieved improvement. Eventually, this resulted in a general improvement of grip strength and decrement in arm disability during follow-up (figures 2A and 2B). Higher and more normal grip strength values were noted compared with grip strength measures at entry (figure 3). A significant association was demonstrated between grip strength and arm disability scale values in these patients (figures 2C and 2D).

The calculated standardised response mean (SRM) scores for grip strength were good (SRM values for the right hands: 1.2, 1.1, 1.3, 1.2, and 0.8; for the left hands: 1.0, 0.8, 1.2, 1.3, and 1.1 at weeks 4, 12, 26, 40, and 52). The patients graded their clinical condition 53 times as

"deteriorating", 38 times as "stable", and 110 times as "improving". These values were significantly associated with the serially obtained grip strength differences in these patients (random effects linear regression analyses: R = 0.50 for association with right hands; R = 0.46 for association with left hands, $p \leq 0.0001$).

Figure 2

Changes of median grip strength for both hands (A) and median arm disability scale scores (B) over time and their association (C & D)



Legend to Figure 2
Grip strength = GS; arm disability scale = ADS. Twenty patients (7 GBS, 13 CIDP) were longitudinally examined. A total of 201 visits were completed. The association between GS and ADS are calculated using "xtreg" (see section statistics) (30). A decrement in ADS corresponds with improvement.

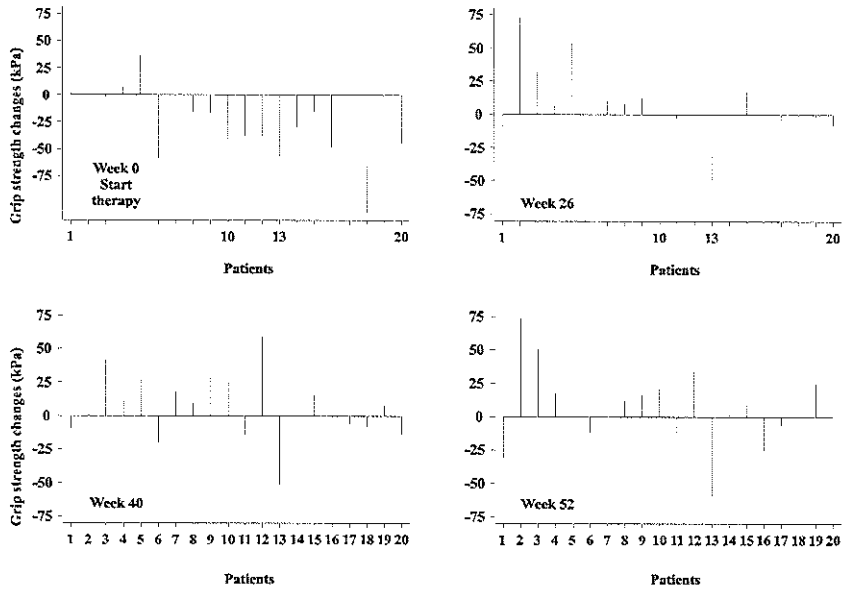
Discussion

The present study provides reference values for grip strength in males and females for both hands using the small bulb in healthy children up to 10 years of age and the medium-sized bulb of the hand-held Vigorimeter in healthy participants 10 years and older. In contrast to what is usually assumed in the literature, these reference values are not straight linearly or quadratic related to age and gender. The relation described in various papers is most probably due to the smaller age-ranges studied, thus possibly acquiring an incomplete view of grip strength in relation to age (9-15). Desrosiers and associates reported on two occasions normative data in healthy individuals age 60 years and older and suggested an association between quadratic-age and grip strength (9,10). In another study, 486 healthy

adults aged 16 to 94 years were examined, but unfortunately there was no stratification for age and gender and no information was provided regarding the healthy state of these adults (11).
To our knowledge, no study has provided reference values for the Vigorimeter in a cohort of healthy controls with an age ranging from 5 to 93 years. An equivalent study using the

Figure 3

Longitudinal grip strength changes in twenty patients with immune-mediated polyneuropathies (right hands)



Legend to figure 3
Each vertical spike represents a measured grip strength value for the right hand of a patient minus the corresponding (age and sex matched) 0.05 quantile grip strength reference value (= grip strength change). It is apparent that most patients improved in time, because most spikes demonstrate a shift toward a (more) positive grip strength change during follow-up. The left hand grip strength changes were quite equivalent and are therefore not shown. Note: nineteen patients completed a one-year follow-up.

Jamar dynamometer was however reported by Mathiowetz (33). The graphical presentation of the reference grip strength values of this device demonstrated approximately the same shape as seen in the current study for both genders (33). Unfortunately, like in many studies, these values were obtained by calculating the means of sequential age-portions in healthy participants (11-15,33). Thus, no consideration was given to the non-Gaussian distribution of the obtained grip strength values, as recommended by Altman for these kind of data (34). In addition to age, multivariate quantile regression analyses demonstrated that hand-circumference was a significant predictor of grip strength in males but not females. The

obtained grip strength values in the current study were consistently higher in men than women. This is not surprising, because anthropometric variables were also constantly higher in men (table 1).

In addition to other reports, the Vigorimeter was demonstrated in the current study to be easily applicable, quick and user friendly with significant correlation to a functional arm scale. Good correlation was also demonstrated between this device and the widely used Jamar dynamometer (9-11). There are however mechanical differences between these two instruments that accounted for the preferential use of the Vigorimeter in the current study. We have previously applied the Jamar dynamometer in a group of patients with polyneuropathies and predominantly distal weakness. These patients consistently reported discomfort when assessing grip strength with this device. Like in the arthritic patients, the Jamar dynamometer caused stress on joints and skin and was less convenient because of its weight and rigidity (11). High reliability values were demonstrated for the Vigorimeter. The reproducibility of grip strength was independent of experience with the Vigorimeter, because the reliability values did not differ between examiners having different degrees of experience. Solgaard *et al.* investigated the accuracy of the Vigorimeter using a "universal testing machine" that compressed the Vigorimeter with known forces and also concluded that it was "very precise" (35).

The usefulness of an outcome instrument not only depends on its simplicity, applicability, validity, and reliability, but also on its ability to be responsive to clinical changes over time (21,22). A heuristic and statistic approach in examining responsiveness of a measure has been proposed (22). Statistical responsiveness captures the ability of an instrument to measure any change, irrespective of its relevance. Heuristic techniques are based on comparing changes as assessed by an outcome measure with an external indicator, for example the clinical-judgement scores by the patient in the current study (21,22). We examined these two approaches in the responsiveness evaluation of grip strength. In the longitudinally examined patients, improvement of grip strength was significantly related to an external criterion, the patients' clinical-judgement-scores, and the standardised response mean (SRM) scores for grip strength were equal to or higher than the proposed value of 0.8 (31,32). However, these SRM values represent responsiveness within a group of patients. It is not clear whether substantial discriminative responsiveness scores will be obtained for the Vigorimeter when evaluating various groups of patients, e.g., in a trial setting comparing a placebo versus a treated group (36).

Grip strength was also significantly associated with the arm disability values over time. These findings imply that, although a simple function in itself, grip strength can be applied as an index of arm functional recovery in patients with immune-mediated polyneuropathies. Hence, grip strength provides information on impairment and indirectly at disability level of outcome.

In conclusion, reference values for grip strength using the hand-held Vigorimeter are presented in a large group of healthy participants. The validity, reliability, and responsiveness of this easily applicable instrument are demonstrated in patients with immune-mediated polyneuropathies. The Vigorimeter is a suitable tool for monitoring treatment efficacy and course of illness in these patients, primarily at the impairment level and secondary as an indicator at the disability level of outcome.

References

1. Medical Research Council. Aids to the investigation of the peripheral nervous system. London: Her Majesty's Stationary Office, 1943:1-2.

2. Kleyweg RP, van der Meché FGA, Schmitz PIM. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré Syndrome. Muscle Nerve 1991;14:1103-1109.

3. Örndahl G, Sellden U, Hallin S, *et al.* Myotonic dystrophy treated with selenium and vitamin E. Acta Med Scand 1986;219:407-414.

4. Rhind VM, Bird HA, Wright V. A comparison of clinical assessments of disease activity in rheumatoid arthritis. Annals Rheum Dis 1980;39:135-137.

5. Sunderland A, Tinson D, Bradley L, *et al.* Arm function after stroke. An evaluation of grip strength as a measure of recovery and a prognostic indicator. J Neurol Neurosurg Psychiatry 1989;52:1267-1272.

6. Toyka KV, Augspach R, Wiethölter H, *et al.* Plasma exchange in chronic inflammatory polyneuropathy: evidence suggestive of a pathogenic humoral factor. Muscle Nerve 1982;5:479-484.

7. Bollag HR, Segmüller G. Zur transmetakarpalen Schrägamputation des Zeigefingers. Schweiz Med Wschr 1975;105:1215-1218.

8. Fünfgeld EW. Vigorimeter: Zur Kraftmessung der Hand und zur Simulationsprüfung. Dtsch Med Wschr 1966;49:2214-2216.

9. Desrosiers J, Bravo G, Hébert R, *et al.* Normative data for grip strength of elderly men and women. Amer J Occup Ther 1995;49:637-644.

10. Desrosiers J, Hébert R, Bravo G, *et al.* Comparison of the Jamar Dynamometer and the Martin Vigorimeter for grip strength measurements in a healthy elderly population. Scand J Rehab Med 1995;27:137-143.

11. Fike ML, Rousseau E. Measurement of adult hand strength: a comparison of two instruments. Occup Ther J Res 1982;2:43-49.

12. Fraser C, Benten J. A study of adult hand strength. Br J Occup Ther 1983;46:296-299.

13. Link L, Lukens S, Bush MA. Spherical grip strength in children 3 to 6 years of age. Amer J Occup Ther 1995;49:318-326.

14. Robertson A, Deitz J. A description of grip strength in pre-school children. Amer J Occup Ther 1988;42:647-652.

15. Thorngren K-G, Werner CO. Normal grip strength. Acta Orthop Scand 1979;50:255-259.

16. Besinger UA, Toyka KV, Hömberg M, *et al.* Myasthenia gravis: Long-term correlation of binding and bungarotoxin blocking antibodies against acetylcholine receptors with changes in disease severity. Neurology 1983;33:1316-1321.

17. Jones E, Hanly JG, Mooney R, *et al.* Strength and function in the normal and rheumatoid hand. J Rheumatol 1991;18:1313-1318.

18. Feinstein AR. Clinimetrics. Yale University Press; New Haven and London, 1987.

19. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. J Neurol Neurosurg Psychiatry 1996;60:127-130.

20. World Health Organisation. International Classification of Impairments, Disabilities, and Handicaps. Geneva: WHO, 1980.

21. Fortin PR, Stucki G, Katz JN. Measuring relevant change: an emerging challenge in rheumatologic clinical trials. Arthritis Rheum 1995;38:1027-1030.

22. Liang MH. Evaluating measurement responsiveness. J Rheumatol 1995;22:1191-1192.

23. Sharrack B, Hughes RAC, Soudain S. Guy's neurological disability scale. Neurology 1996;243(suppl2):S32.

24. Ad Hoc subcommittee of the American Academy of Neurology AIDS task force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

25. Asbury AK, Cornblath DR. Assessment of current diagnostic criteria for Guillain-Barré syndrome. Ann Neurol 1990;27(suppl):S21-S24.

26. Miescher GC, Steck AJ. Paraproteinaemic neuropathies. Baillieres Clin Neurol 1996;5:219-232.

27. American society of hand therapists. Clinical assessment recommendations, 2nd ed, Chicago, Author, 1992.

28. Gould WW, Rogers WH. Quantile regression as an alternative to robust regression. Proceedings of the statistical computing section. Alexandria, VA: American Statistical Association, 1994.

29. Herndon JE II, Harrell FE Jr. The restricted cubic spline hazard model. Communications in Statistics-Theory and Methods. 1990;19:639-663.

30. Dwyer J, Feinleib M. Introduction to statistical models for longitudinal observation. In: Dwyer J, Feinleib M, Lippert P, Hoffmeister H, eds. Statistical models for longitudinal studies of health. New York, Oxford University Press, 1992: 3-48.

31. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;28:632-642.

32. Cohen J. Statistical power analysis for the behavioural sciences. New York: Academic Press, 1977.

33. Mathiowetz V. Grip and pinch strength measurements. In: Amundsen LR, ed. Muscle strength testing. Instrumented and non-instrumented systems. New York: Churchill Livingstone, 1990:163-177.

34. Altman DG. Practical statistics for medical research. 6th ed. London: Chapman & Hall, 1995:203-205, 213-216.

35. Solgaard S, Kristiansen B, Jensen JS. Evaluating of instruments for measuring grip strength. Acta Orthop Scand 1984;55:569-572.

36. Guyatt G, Walter S, Normal G. Measuring change over time: Assessing the usefulness of evaluative instruments. J Chron Dis 1987;40:171-178.

Addendum

HET WITTE WEEKBLAD EDITIE 54-55 WEEK 48/266/1998

Spierzwakte door zenuwaandoeningen

Knijpkrachtonderzoek geeft informatie over ziekteverloop

Haarlemmermeer – Zenuwaandoeningen kunnen leiden tot zwakte van verschillende spieren. Het meten van de knijpkracht van een patiënt kan informatie geven over toenemende zwakte of over het effect van een bepaalde behandeling. Dr. Merkies uit Nieuw-Vennep doet onderzoek naar de 'knijpkracht' van gezonde mensen om zo eerst vast te stellen wat de normale knijpkracht is. Hij zoekt daarvoor vrijwilligers.

Dr. Merkies werkt op de afdeling neurologie van het Academisch Ziekenhuis in Rotterdam. Hij wil de knijpkracht van in totaal vijf-Honderd gezonde vrijwilligers in de leeftijd van 5 jaar tot boven de 80 jaar meten. Het meten van de knijpkracht gebeurt met behulp van een rubber balletje. Voordat de knijpkracht van een vrijwilliger wordt gemeten, worden lengte, gewicht en hand-omtrek vastgesteld.

Tot dusver heeft dr. Merkies tweehonderd mensen onderzocht, overwegend uit Nieuw-Vennep. Hij heeft onder ander het Herbert Vissers College bezocht en het Op Dreef basisschool. Voldoende jongeren tot 18 jaar hebben een bijdrage aan dit onderzoek geleverd. Ook heeft hij een aantal zelfstandig wonende senioren in Nieuw-Vennep thuis bezocht bij wie het onderzoek in het algemeen niet langer duurde dan vijf minuten. Het onderzoek is pijnvrij en slechts de handen en de polsen zullen ontbloot worden. De bevindingen zullen anoniem geregistreerd worden en verder gebruikt om de gemiddelde knijpkracht per leeftijd en geslacht te berekenen. Deze waardes zullen in de praktijk gebruikt worden om te kunnen aangeven of een patiënt met een bepaalde zenuwaandoening een verminderde knijpkracht heeft of niet.

Op de afdeling neurologie van het Academisch Ziekenhuis in Rotterdam wordt veel onderzoek gedaan naar zenuwaandoeningen die tot zwakte van verschillende spieren kunnen leiden. Deze onderzoeken vinden plaats onder leiding van prof. dr. F.G.A. van der Meché, hoofd van de afdeling neurologie. Een manier om spierzwakte in kaart te brengen is door het meten van de knijpkracht bij patiënten. De indruk bestaat dat het vaststellen van de knijpkracht ook veel zegt over bijvoorbeeld het effect van een bepaalde behandeling van een patiënt. Door de knijpkracht van een patiënt in de tijd te volgen, kan worden vastgesteld of de kracht vermindert of niet. Voor men echter kan zeggen of een patiënt een verminderde knijpkracht heeft, is het noodzakelijk eerst bij gezonde mannen en vrouwen na te gaan wat de normale knijpkracht is bij een bepaalde leeftijd. Wie 18 jaar of ouder is, gezond, en een bijdrage wil leveren aan dit onderzoek, kan vrijblijvend in contact treden met dokter Merkies, telefoon 0252 686287. Dat nummer kan ook gebeld worden als men meer informatie wil.

Chapter 6

Clinimetric evaluation of an overall disability scale in immune-mediated polyneuropathies

Merkies I.S.J.[1], MD, Schmitz P.I.M.[2], PhD, van der Meché F.G.A.[1], PhD, Samijn J.P.A.[1], MD, van Doorn P.A.[1], PhD, for the Inflammatory Neuropathy Cause and Treatment (INCAT) group[3]

From the [1]Department of Neurology, [2]Department of Statistics-Daniel den Hoed Cancer Centre, University Hospital Rotterdam/Erasmus University, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands. [3]See Chapter 2 for INCAT-members.

Submitted for publication

Abstract

Objectives: To determine the validity, reliability, and responsiveness of an overall disability sumscore in immune-mediated polyneuropathies.
Methods: Three impairment measures (MRC sumscore, sensory sumscore, grip strength with the Vigorimeter) and three disability scales (an overall disability sumscore [ODSS], Hughes' functional grading scale [f-score], Rankin scale) were assessed in 113 clinically stable patients (83 who experienced Guillain-Barré syndrome (GBS), 22 with chronic inflammatory demyelinating polyneuropathy (CIDP), 8 patients with a gammopathy related polyneuropathy). The ODSS was also utilised serially in twenty patients with recently diagnosed GBS (n=7) or CIDP (n=13) and changing clinical conditions. The ODSS comprises a functional description of arm and leg functions in a checklist form, suitable for interviewing patients. The arm component addresses daily arm activities and the leg component highlights problems regarding walking. The validity of the ODSS was examined by correlation and regression studies with the other measures and its reliability was investigated by calculating the inter-/intra-observer reliability. Its responsiveness was analysed using the standardised-responsive-mean (SRM) technique. Multiple regression studies were performed to compare the impact of impairment disturbances (independent variables) on the various disability scales separately (dependent variable).
Results: Good validity (stable group: Spearman's Rank test: r=-0.45 to -0.74 and r=0.41 to 0.79; longitudinal group: Intraclass correlation coefficient: R=0.69-0.89; $p<0.006$ for all associations) and reliability (R=0.90-0.95;$p<0.0001$) were demonstrated for the ODSS. The SRM values for the ODSS were high (>0.8), indicating good responsiveness. Impairment measures accounted for a higher variance-proportion of the ODSS compared with the f-score and Rankin (R^2=0.64 versus 0.56 and 0.45, respectively).
Conclusion: All clinimetric requirements were demonstrated for the overall (arm and leg) disability sumscore in immune-mediated polyneuropathies. Also, this scale adequately monitored impairment disturbances leading to disability. Its use is therefore suggested in the evaluation of immune-mediated polyneuropathies.

Introduction

Clinical assessment in patients with Guillain-Barré syndrome (GBS), chronic inflammatory demyelinating polyneuropathy (CIDP), and polyneuropathy associated with a monoclonal gammopathy of undetermined significance (MGUSP) has been traditionally focused on impairment and less frequently on disability (1,2). The most commonly used impairment scales are motor scales based on the Medical Research Council (MRC) grading system and different sensory outcome measures including various sensation modalities (3). Disability, on the other hand, has been primarily evaluated using the (modified) Hughes' functional grading scale (f-score) and the (modified) Rankin scale (4-7). Regarding the f-score, its clinimetric properties have been demonstrated in patients with GBS (7). Contrary, no formal clinimetric evaluation of the Rankin scale has been performed in patients with polyneuropathies. Despite the simplicity and obvious face validity, the clinical use of the Hughes' scale and Rankin scale is somewhat limited, since their emphasis is strongly directed towards mobility, thus not providing information regarding the arms.
To date, only a limited number of papers have also evaluated specific arm disability entities in patients with immune-mediated polyneuropathies (8-12). In a recent study, a disability scale was applied in these conditions that described adequately functional disturbances of the legs and arms (11). This overall disability sumscore (ODSS), as part of the Guy's neurological disability scale, demonstrated a significant correlation with the "*INCAT*" sensory sumscore, hereby fulfilling its validity (11,13; see appendix II, page 210). However, further clinimetric evaluation of the ODSS is required before its general use in patients with immune-mediated polyneuropathies (14). Prompted by these observations, we investigated the validity, reliability, and responsiveness of the ODSS. Moreover, this scale was compared with the f-score and Rankin scale to determine which of these disability measures would have the strongest association with a group of impairment disturbances.

Patients and Methods

Patients
One hundred and thirteen patients (83 GBS, 22 CIDP, 8 MGUSP) with a stable clinical condition were recruited from the Rotterdam immune-mediated polyneuropathy databank and the Dutch GBS study group (stable group). Patients with GBS, CIDP or MGUSP were recruited, since it was argued that these disorders represent parts of a continuum regarding their neuromuscular dysfunction pattern (2). The selected patients still had residual symptoms or signs due to their illness, representing a broad range of disability. Nine CIDP patients required interval treatment ranging from weeks to months, with intravenous immunoglobulin (IVIg). With this therapy their clinical condition has been stable for more than 6 months. Six patients with MGUSP (three with IgG, two with IgM, and one patient with IgG+IgM) had a demyelinating polyneuropathy with minor concurrent axonal damage in three. An axonal polyneuropathy was diagnosed in the remaining two patients with MGUSP (one IgA and one IgG type).
Twenty consecutive patients with recently diagnosed sensory-motor GBS (n = 7) or CIDP (n = 13) and changing clinical conditions were enrolled to investigate the responsiveness of the ODSS (longitudinal group). All GBS and CIDP patients met the international criteria for

their illness (15,16). The diagnosis MGUSP was established after excluding all possible causes for the gammopathy and polyneuropathy (17).

Assessment tools/scales

The *MRC sumscore* is a summation of the Medical Research Council grades (range: 0 - 5) given in full numbers of the following muscle pairs: upper arm abductors, elbow flexors, wrist extensors, hip flexors, knee extensors, and foot dorsal flexors (7). The MRC sumscore ranges from 0 ("total paralysis") to 60 ("normal strength"). Good validity and inter-observer reliability were provided for this scale in patients with GBS (7).

The *"INCAT" sensory sumscore* was recently introduced and extensively evaluated in terms of its clinimetric soundness in patients with immune-mediated polyneuropathies (11). In brief, this sensory scale comprises pin-prick and vibration sense plus a two-point discrimination value in the arms and legs and ranges from 0 ("normal sensation") to 20 ("most severe sensory deficit") (11).

The *Vigorimeter* (VM) (Martin, Tuttlingen, Germany) is a hand-held dynamometer to measure grip strength (8). Good clinimetric properties were demonstrated in patients with immune-mediated polyneuropathies (8).

The *Overall Disability sumscore (ODSS)* is composed by a recently published arm and leg disability scale with a total score ranging from 0 ("no signs of disability") to 12 ("most severe disability score") (11,13; see appendix II, page 210). The ODSS comprises a good functional description of the arms and legs in a checklist form suitable for interviewing patients. Daily arm activities like dressing upper part of the body, doing and undoing buttons and zips, washing and brushing hair, using a knife and fork and turning a key in a lock are scored as being "not affected", "affected but not prevented" or "prevented". Subsequently, these results are translated into an arm grade (score range: 0 [normal arm abilities] to 5 [severe symptoms and signs in both arms preventing all purposeful movements]). The leg scale highlights problems regarding walking taking into account the use of a device. The results are also translated into a leg grade (score range: 0 [walking is not affected] to 7 [restricted to wheelchair or bed most of the day, preventing all purposeful movements of the legs]) (13). The selected arm and leg disability scales are subsets of a more comprehensive Guy's neurological disability scale (13). Good clinimetric requirements have been recently demonstrated for all components of the Guy's scale in patients with multiple sclerosis (13).

The Modified *Hughes' functional grading scale* (*f-score*) assesses the functional ability of the patients with a strong emphasis on mobility (7). The f-score of the patients included in this study ranged from 0 (no symptoms or signs) to 5 (requiring artificial ventilation for at least part of the day) (7).

The *Rankin scale* has been primarily used in patients with stroke (6). The grades of this scale range from: 0 (no symptoms at all) to 5 (severe disability, bedridden, incontinent, and requiring constant nursing care and attention) (6). No formal clinimetric evaluation has been performed in patients with immune-mediated polyneuropathies.

Test procedures

General aspects. All participants gave informed consent before the study. All measures were obtained in a quiet and comfortably warm room at our outpatient clinic. The assessments were performed in a random order. For the assessment of general strength, a standardised joint and limb position as well as the point at which counter-force was administered was

defined before the start of the study and taken at examination of each muscle group (page 205). Sensory modalities were examined in triplicate under the earlier prescribed standard conditions with the patients lying in supine position (11). For the assessment of grip strength with the VM, all patients were examined according to the standard conditions described before (8,18).

The study took place between December 1998 and January 2000 and was performed on behalf of the "Inflammatory Neuropathy Cause And Treatment (*INCAT*) group", a collaborating force of European senior neurologists with special interest in neuro-immunological illnesses.

Validity and Reliability. Validity of the ODSS scale was investigated by correlation and regression studies with the other outcome measures. For the assessment of reliability and construct validity of the selected scales in the stable group of 113 patients, two neurologists and six experienced residents in neurology formed 28 different couples. Preceding the study, all investigators received instructions in assessing the outcome measures. Twenty-seven ("variable") couples investigated a total of 68 patients (2-3 patients for each couple). The "experienced" couple (IM+JS) examined the remaining 45 stable patients. The latter couple was formed to examine the effect of training (and thus a possible increase in reliability) when using the ODSS often.

The patients were examined on two different occasions at our outpatient clinic. During the first visit the two members of an appointed pair performed their scores independently and consecutively (usually within 2 hours) (inter-observer measures). Within 2-4 weeks, the patient returned for a second visit and only one investigator of the earlier assigned pair examined the patient again (intra-observer values) without having access to previous results. The assessments sequence at entry and the examination at the second visit were equally distributed among the members of an assigned couple. Eventually, each member of a couple examined the approximately same number of patients. With the exception of the f-score, all scales were assessed at each visit in all patients. For the validity and the regression model studies, only the recruited scales' values at one visit were used.

Responsiveness. Twenty consecutive patients were longitudinally examined by the same clinician (IM). The ODSS was assessed at study entry and at the weeks 2, 4, 8, 12, 16, 21, 26, 32, 40, and 52 of follow-up with additional clinical investigations if necessary.

Statistics

Validity and Reliability. In the stable group, correlation between ODSS and the other outcome measures was analysed using Spearman's Rank correlation test. Also, random effects linear regression analyses were performed of the ODSS on the other scales in the longitudinal group, taking into account the association of the data caused by the longitudinal structure. The latter was achieved using the program "xtreg" in STATA 6.0 (StataCorp. 1997. Stata Statistical Software: Release 6.0. College Station, TX), which is based upon a cross-sectional time-series regression model as described by Dwyer and Feinleib (19). The inter- and intra-rater reliability of the ODSS was quantified by estimation of the intraclass correlation coefficient using a one-way random effects analysis-of-variance model for the two investigator ("experienced" and "variable") groups.

Responsiveness. Responsiveness was investigated by calculating the standardised response mean (SRM) scores for the ODSS at various arbitrarily chosen occasions during follow-up (weeks 12, 26, 40, 52) (20). SRM is equal to the mean change in score divided by the

standard deviation of the change in score (SRM = μi - μo/SD(μi - μo); μi = mean ODSS value of the longitudinally examined group at week = i; μo = mean ODSS value at week = 0 [entry]) (20). A value between 0.5 and 0.8 is considered moderate, and 0.8 or greater as high responsiveness (20,21). All analyses were performed using Stata 6.0 for Windows 95. A value of $p < 0.05$ was considered statistically significant.

Comparative study. In the stable group, uni- and multivariate linear regression analyses were performed to determine which disability measure (ODSS, f-score, or Rankin scale; dependent variable) had the strongest association with a group of impairment measures (MRC sumscore, "*INCAT*" Sensory sumscore, Grip strength by the Vigorimeter; independent group). If necessary, a transformation of the dependent variable (for example by logarithmic conversion) was performed to obtain a normal distribution. Univariate regression studies were primarily performed, striving for the best fit between the dependent and independent variable. This was achieved through systematic evaluation of constructed graphs with linear regression studies that included a restricted cubic spline function on the independent variable (22). Subsequently, multivariate linear regression analyses were performed, with a backward stepwise eliminating strategy to construct the final models.

Only the results that included the right hand grip strength values will be presented, since these findings turn out to be similar to the regressions that incorporated the left hand values.

Table 1

Basic characteristics of patients with immune-mediated polyneuropathies

Stable group of patients (n = 113; GBS 83, CIDP 22, MGUSP 8)	
Median age at start of the study (range) [years]	56 (14 – 84)
Median MRC sumscore (score range: 0 – 60)	54 (18 – 60)
Median "*INCAT*" Sensory sumscore (score range: 0 – 20)	3 (0 – 18)
Median grip strength values with the Vigorimeter (score range: 0 – 160 kPa)	
Right hands	65 (0 – 158)
Left hands	62 (0 – 160)
Median Overall Disability sumscore (score range: 0 – 12)	
At entry	4 (0 – 11)
Second visit	4 (0 – 12)
Third visit	3 (0 – 12)
Median f-score (score range: 0 - 5)	2 (1 – 4)
Median Rankin score (score range: 0 – 5)	2 (0 – 4)
Longitudinal group of patients (n = 20; GBS 7, CIDP 13)	
Median age at start of the study (range) [years]	54 (15 – 70)
Median Overall Disability sumscore (score range: 0 – 12)	
At entry	5 (3 – 11)
At 12 weeks of follow-up	3 (0 – 10.5)
At 26 weeks of follow-up	2.5 (0 – 9.5)
At 40 weeks of follow-up	2 (0 – 9)
At 52 weeks of follow-up	2 (0 – 9)

Legend to Table 1

GBS = Guillain-Barré syndrome; CIDP = chronic inflammatory demyelinating polyneuropathy; MGUSP = polyneuropathy associated with a monoclonal gammopathy of undetermined significance. *INCAT* = inflammatory neuropathy cause and treatment group.

Results

General aspects. All eight examiners who investigated the patients concluded that the selected ODSS was easily applicable and required less than two minutes for completion. The stable group of patients (54 females and 59 males) had a median duration of symptoms till

onset of the study of 5.1 years. Seven of these patients were bed-bound and fourteen required assistance or a device to walk short distances. The remaining 92 patients could walk independently. The corresponding median values and ranges for all scales in these patients are presented in table 1.

Validity, inter- and intra-observer Reliability. The correlation studies between the ODSS and other scales and reliability values of the ODSS in the stable group of patients are presented in Table 2. Significant validity and good reliability were demonstrated for the ODSS by the "experienced" and "variable" couples of investigators (Table 2). In the longitudinal group, significant associations were also obtained between the ODSS and other measures (ODSS versus MRC sumscore: Random effects analysis-of-variance, R=0.89; versus Sensory sumscore: R=0.74; versus grip strength: R=0.72 [right hands] and R=0.69 [left hands]; versus f-score: R=0.86; versus Rankin: R=0.88; p<0.0001 for all associations).

Table 2

Validity and reliability analyses of the Overall disability sumscore (ODSS) in a stable group of patients with immune-mediated polyneuropathies (n=113)

	"Experienced" couple of examiners (couple number 1; 45 patients)		"Variable" couples of examiners (couples number 2-28; 68 patients)	
Validity	Spearman rank correlation coefficient (r)	p-value	Spearman rank correlation coefficient (r)	p-value
Overall Disability sumscore versus				
MRC sumscore	-0.45	0.002	-0.71	<0.0001
"*INCAT*" sensory sumscore	0.41	0.006	0.56	<0.0001
Grip strength by the Vigorimeter				
Right hands	-0.54	0.0002	-0.70	<0.0001
Left hands	-0.53	0.0002	-0.74	<0.0001
F-score	0.78	<0.0001	0.74	<0.0001
Rankin	0.78	<0.0001	0.79	<0.0001
Reliability	Intraclass correlation coefficient (R)	p-value	Intraclass correlation coefficient (R)	p-value
Overall Disability sumscore				
Inter-observer agreements	0.95	<0.0001	0.90	<0.0001
Intra-observer agreements	0.95	<0.0001	0.93	<0.0001

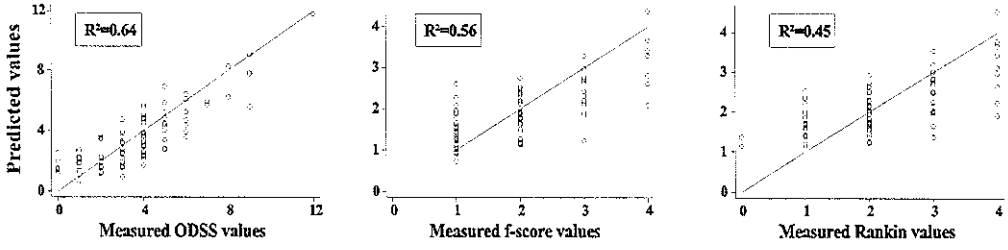
Responsiveness. Eight females and twelve males were examined longitudinally. At study entry, four patients were bed-bound, one requiring artificial ventilation, and nine patients were unable to walk independently. All patients experienced general loss of strength, sensory disturbances, and deficit in daily functional activities. Two hundred and one visits were completed during a follow-up period of 40 to 58 (median: 52) weeks. Nineteen patients completed a one-year follow-up. With the exception of one GBS patient who only experienced mild symptoms, all patients received initial treatment with IVIg (0.4 grams/kilogram body weight/day for 5 consecutive days). All but one patient with CIDP showed good functional improvement on IVIg during follow-up. The non-responder received a treatment course of oral prednisone and also demonstrated improvement. The GBS patients did not show any deterioration during follow-up and improved gradually over time. After initial improvement, all 12 IVIg responsive CIDP patients needed interval therapy with IVIg (1-2 days 0.4 grams/kg/day; intervals: 3 - 21 weeks) to maintain earlier achieved improvement. Eventually, all patients demonstrated during follow-up a general decrement in impairments and improvement of functional abilities. Improvement in the longitudinal group resulted in a general reduction in the ODSS values (indicating

improvement) (median values: 3, 2.5, 2, 2 at weeks 12, 26, 40, and 52, respectively) compared with the median entry value of 5 (Wilcoxon signed-rank test: $p \leq 0.0008$ for all comparisons). Good SRM scores were obtained for the ODSS in these patients (SRM values: 1.2, 1.5, 1.4, and 1.4 at weeks 12, 26, 40, and 52, respectively).

Comparative study. The MRC sumscore was the strongest predictor of disability compared with the grip strength (Vigorimeter) and sensory sumscore (univariate linear regression analyses; on ODSS: MRC sumscore, R^2=0.45; grip strength, R^2=0.40; sensory sumscore, R^2=0.21; regressions on the f-score: MRC sumscore, R^2=0.43; grip strength, R^2=0.34; sensory sumscore, R^2=0.16; regressions on the Rankin: MRC sumscore, R^2=0.34; grip strength, R^2=0.24; sensory sumscore, R^2=0.14). Overall, a higher proportion of variance in disability, explained by the impairment measures, was captured by the ODSS compared with the f-score and Rankin (Figure).

Figure

Patients' level of disability explained by impairment variables (MRC sumscore, Vigorimeter, Sensory sumscore) in immune-mediated polyneuropathies



Legend to Figure

In the stable group (n=113), multivariate regression analyses were performed of the disability scales separately (overall disability sumscore [ODSS], Hughes' functional grading scale [f-score], or Rankin scale; dependent variable) on the group of impairment measures (MRC sumscore, "*INCAT*" Sensory sumscore, and grip strength by the Vigorimeter; independent group). The "Predicted values" were obtained from these regressions. Only the results that included the right hand grip strength values were presented, since these findings turn out to be similar to the regressions that incorporated the left hand values.

Discussion

In the current study, the clinimetric requirements like being easily applicable, valid, reliable, and responsive to clinical changes over time are demonstrated for the overall disability scale (ODSS) in patients with immune-mediated sensory-motor polyneuropathies (14). As stated earlier, this scale highlights not only problems regarding walking, but also addresses daily arm activities. In these conditions, its concept is therefore more comprehensive than the widely used Hughes' functional grading scale (f-score) and Rankin scale, who are mainly directed towards mobility, not providing information on the arms (4,6). Also, general loss of strength and sensory deficit leading to disability was better monitored by the ODSS

compared with the f-score and Rankin scale. Hence, the preference use of the ODSS is suggested for evaluation of outcome at the disability level in these conditions.
The impairment variables explained two third of the disability, as registered by the ODSS. This finding implicates that other clinical impairment variables should be considered in future studies as possible contributors to disability. Variables like general fatigue and depression have been advocated as important events in patients with immune-mediated polyneuropathies that may lead to functional deficit (23-25). Lennon and associates reported six reasons for persistent disability in patients with Guillain-Barré syndrome (25). These were: muscle weakness, sensory dysfunction, contractures, fatigue, other medical conditions, and psychological factors like anxiety, depression, and motivation (25).
In the current study, weakness, as measured with the MRC sumscore, was the most important independent explanatory factor of the patients' level of disability. This finding is consistent with a recent paper, addressing outcome in various forms of polyneuropathy (26).
With respect to the aims of the current study, some methodological issues should be addressed. First, the obtained SRM scores for the ODSS only demonstrated within-group responsiveness. It is not clear whether substantial discriminative responsiveness scores will be obtained for the ODSS when evaluating various groups of patients, e.g. in a trial setting comparing a placebo versus a treated group (27). Second, uni- and multivariate linear regression analyses of the f-score and Rankin scale were performed, despite the fact that these outcome measures are ordinal constructed. An ordinal logit estimation model, as described by the program "ologit" in Stata 6.0, was also applied on these ordinal variables, but since the description of these analyses were rather complex and the results quite similar to the linear regression studies, we decided to present the data as such to strive for clarity.
In conclusion, the simplicity, validity, reliability, and responsiveness are demonstrated for the overall (arm plus leg) disability sumscore in patients with immune-mediated polyneuropathies. Also, impairment leading to disability was better monitored by the overall disability sumscore compared with other tested disability measures. Hence, the use of the overall disability sumscore is suggested for monitoring outcome at the disability level in immune-mediated polyneuropathies.

References

1. World Health Organisation. International classification of impairments, disabilities, and handicaps. Geneva: WHO, 1980.

2. Hartung H-P, van der Meché FGA, Pollard JD. Guillain-Barré syndrome, CIDP and other chronic immune-mediated neuropathies. Curr Opin Neurol 1998;11:497-513.

3. Medical Research Council. Aids to the investigation of the peripheral nervous system. London: Her Majesty's Stationary Office, 1943.

4. Hughes RAC, Newsom-Davis JM, Perkin GD, *et al.* Controlled trial prednisolone in acute polyneuropathy. Lancet 1978;2:750-753.

5. Van der Meché FGA, Schmitz PIM, and the Dutch Guillain-Barré study group. A randomised trial comparing intravenous immune globulin and plasma exchange in Guillain-Barré syndrome. N Engl J Med 1992;326:1123-1129.

6. van Swieten JC, Koudstaal PJ, Visser MC, *et al.* Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988;19:604-607.

7. Kleyweg RP, van der Meché FGA, Schmitz PIM. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré syndrome. Muscle Nerve 1991;14:1103-1109.

8. Merkies ISJ, Schmitz PIM, Samijn JP, *et al.* Assessing grip strength in healthy individuals and patients with immune-mediated polyneuropathies. Muscle Nerve 2000;23:1393-1401.

9. Plasma exchange/Sandoglobulin Guillain-Barré syndrome trial group. Randomised trial of plasma exchange, intravenous immunoglobulin, and combined treatments in Guillain-Barré syndrome. Lancet 1997;349:225-230.

10. Rees JH, Soudain SE, Gregson NA, *et al.* Campylobacter jejuni infection and Guillain-Barré syndrome. N Engl J Med 1995;333:1374-1379.

11. Merkies ISJ, Schmitz PIM, van der Meché FGA, *et al.* Psychometric evaluation of a new sensory scale in immune-mediated polyneuropathies Neurology 2000;54:943-947.

12. Guillain-Barré syndrome Steroid Trial group. Double-blind of intravenous methylprednisolone in Guillain-Barré syndrome. Lancet 1993;341:586-590.

13. Sharrack B, Hughes RAC, Soudain S. Guy's neurological disability scale [abstract]. Neurology 1996;243(suppl 2):S32.

14. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. J Neurol Neurosurg Psychiatry 1996;60:127-130.

15. Asbury AK and Cornblath DR. Assessment of current diagnostic criteria for Guillain-Barré Syndrome. Ann Neurol 1990;27(suppl):S21-S24.

16. Ad Hoc subcommittee of the American Academy of Neurology AIDS task force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

17. Miescher GC, Steck AJ. Paraproteinaemic neuropathies. Baillieres Clin Neurol 1996;5:219-232.

18. American society of hand therapists. Clinical assessment recommendations, Chicago: 2nd ed, 1992.

19. Dwyer J, Feinleib M. Introduction to statistical models for longitudinal observation. In: Dwyer J, Feinleib M, Lippert P, Hoffmeister H, editors. Statistical models for longitudinal studies of health. New York: Oxford University Press, 1992; 3-48.

20. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990;28:632-42.

21. Cohen J. Statistical power analysis for the behavioural sciences. New York: Academic Press; 1977.

22. Herndon JE II, Harrell FE Jr. The restricted cubic spline hazard model. Comm Stat Theory Methods 1990;19:639-663.

23. Eisendrath SJ, Matthay MA, Dunker JA, *et al.* Guillain-Barré syndrome: Psychosocial aspects of management. Psychosomatics 1983;24:465-475.

24. Merkies ISJ, Schmitz PIM, Samijn JPA, *et al.* Fatigue in immune-mediated polyneuropathies. Neurology 1999;53:1648-1654.

25. Lennon SM, Koblar S, Hughes RAC, *et al.* Reasons for persistent disability in Guillain-Barré syndrome. Clin Rehabil 1993;7:1-8.

26. Molenaar DS, Vermeulen M, de Visser M, *et al.* Impact of neurologic signs and symptoms on functional status in peripheral neuropathies. Neurology 1999;52:151-156.

27. Guyatt G, Walter S, Norman G. Measuring change over time: Assessing the usefulness of evaluative instruments. J Chron Dis 1987;40:171-178.

Chapter 7

Chronic motor neuropathies: Treatment with Interferon-β1a after failure of conventional therapies

Merkies ISJ[1], van Doorn PA[1], Schmitz PIM[2], Meulstee J[1], van der Meché FGA[1], for the Inflammatory Neuropathy Cause And Treatment (INCAT) group[3]

From the Department of [1]Neurology, [2]Department of Statistics - Daniel den Hoed Cancer Centre, Erasmus University Medical Centre Rotterdam, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands. [3]See Chapter 2 for INCAT-members.

J Neurol Neurosurg Psychiatry, 1999;66:197-201; published with permission from the BMJ Publishing Group

published under the former name Martina ISJ

Abstract

Objectives: The effect of Interferon-ß1a (Rebif®) was studied in patients with chronic motor neuropathies not improving after conventional treatments like immunoglobulins, steroids, cyclophosphamide or plasma exchange.
Methods: A prospective open study was performed with duration of 6 to 12 months. Three patients with a multifocal motor neuropathy and one patient with a pure motor form of chronic inflammatory demyelinating polyneuropathy were enrolled. Three patients had anti-GM1 antibodies. Treatment consisted of subcutaneous injections of Rebif (6 MIU), three times a week. Primary outcome was assessed at the disability level using the Nine-hole peg, the Ten-metre walking test, and the modified Rankin scale. Secondary outcome was measured at the impairment level using a slightly modified MRC Sumscore.
Results: All patients showed a significant improvement on the modified MRC sumscore. The time required to walk 10 meters and to fulfil the Nine-hole peg test was also significantly reduced in the first three months in most patients. However, the translation of these results to functional improvement on the modified Rankin scale was only seen in two patients. There were no severe adverse events. Motor conduction blocks were partially restored in one patient only. Anti-GM1 antibody titres did not change.
Conclusion: These findings indicate that severely affected patients with chronic motor neuropathies not responding to conventional therapies may improve when treated with interferon-ß1a. From this study it is suggested that interferon-ß1a should be administered in patients with chronic motor neuropathies for a period up to 3 months before deciding to cease treatment. A controlled trial is necessary to confirm these findings.

Introduction

Multifocal motor neuropathy is a chronic immune-mediated demyelinating neuropathy (1,2). Patients with multifocal motor neuropathy mostly have a stepwise progression of asymmetrical muscle weakness and amyotrophy localised in the anatomical distribution areas of peripheral nerves. Sensory symptoms are generally not present. The electrophysiological hallmark of multifocal motor neuropathy is persistent conduction block. Most patients with multifocal motor neuropathy have high titre antibodies against the ganglioside GM1 (1,2). Clinically, multifocal motor neuropathy is also described as an asymmetrical pure motor variant of chronic inflammatory demyelinating polyneuropathy with multifocal motor conduction blocks. Especially, during the evolution of multifocal motor neuropathy the multifocal character may gradually evolve in a more or less symmetrical pattern, clinically resembling the motor form of chronic inflammatory demyelinating polyneuropathy. Pathological studies have also linked multifocal motor neuropathy to chronic inflammatory demyelinating polyneuropathy (3,4).
The first line of treatment of patients with multifocal motor neuropathy constitutes high dosage of intravenous immunoglobulines (IVIg) (5,6). Initial reports suggested benefit from treatment with cyclophosphamide (1,7). However, not all patients do improve after these treatments. At least in the group of non-responders, there is a need for new treatment modalities. Such a new treatment might be interferon-β1a (IFN-β1a).
The main objective of this study was to investigate whether treatment with IFN-β1a results firstly in improvement at the level of disability and secondary at the impairment level in severely affected patients with chronic motor neuropathies not improving after conventional therapies such as IVIg, cyclophosphamide, steroids, or plasma exchange. Additionally, the influence of IFN-β1a on neurophysiological findings and anti-GM1 antibody titres was investigated.

Patients and methods

Patients' characteristics (table) and therapy
Four patients entered this prospective open study. Three patients were diagnosed as having multifocal motor neuropathy based on the clinical and electrophysiological characteristics for multifocal motor neuropathy (2). The fourth patient had a chronic symmetrical pure motor neuropathy with rapidly evolving symptoms at onset in the lower limbs. This patient, met the clinical, electrophysiological, and CSF criteria for the diagnosis chronic inflammatory demyelinating polyneuropathy (8). Immunofixation in this patient, however, showed an IgM-lambda and an IgG-kappa monoclonal gammopathy. Extended haematological investigation showed no other abnormalities.
Two patients with multifocal motor neuropathy and the patient with chronic inflammatory demyelinating polyneuropathy had IgM anti-GM1 antibodies. Anti-MAG (myelin associated glycoprotein) antibodies were absent in the patient with chronic inflammatory demyelinating polyneuropathy. The duration of the symptoms ranged from 7 to 9 years before the start of Interferon-β1a treatment. All patients developed marked amyotrophy and their ambulation also decreased gradually during the disease period. At entry to this study, one patient was almost always wheelchair dependent and two patients needed a walking stick and ankle

Table

Clinical data before onset of treatment with Interferon-β1a (Rebif)

Patient	Age/ Gender	Diagnosis	Onset of symptoms	Initially affected motor nerves	Duration of illness	Previous treatments	Anti-GM1 antibodies
1	39/M	MMN	UL/D/A	L Radial/L Ulnar R Radial	9 years	IVIg[1]/Cyclophos (2x) [2a]/Predn[3] Cyclophos[2b]+Predn[3]/IVIg[1]+Mpredn[4]	+ (1:12.800)
2	60/F	MMN	UL/D/A	R Median/L Ulnar	8 years	IVIg (3x) [1]	-
3	53/F	MMN	UL/D/A	L Ulnar	7 years	IVIg (2x)[1]/ IVIg[1]+ Mpredn[4]	+ (1:200)
4	54/M	CIDP	LL/D>P/S	L+R Posterior tibial and Peroneal	9 years	IVIg (2x)[1]/IVIg[1]+Mpredn[4]/Predn[3]/PE[5] Cyclophos[2a]+Predn[3]/Cyclophos[2c]	+ (1:12.800)

Legends to Table

M=male; F=female; MMN=multifocal motor neuropathy; CIDP=chronic inflammatory demyelinating polyneuropathy; UL=upper limb; LL=lower limb; D=distal; P=proximal; A=asymmetrical; S=symmetrical; L=left; R=right; IVIg=intravenous immunoglobulin; Cyclophos=cyclophosphamide; (M)Predn=(Methyl)prednisolone; PE=plasma exchange; [1]=0.4g/kg body weight/day for 5 days; [2a]=0.5g/day intravenous for 14 days; [2b]=0.15g/day orally for half a year; [2c]=monthly 0.5g/day orally for 4 days, for half a year; [3]=60-80mg/day orally for 6 weeks, thereafter tapering to zero in half a year; [4]=0.5g/day for 5 days; [5]=2 exchange sessions (each 2.5 litres plasma)/week for 5 consecutive weeks.

orthesis at both sites to cover short distances and a wheelchair for longer distances. Patient number 3 did not experience many problems when walking very short distances, but she noticed that her legs gave way after walking for 5 to 10 minutes. Her walking endurance was also deteriorating and she could walk outdoors for only 15 minutes. The patients received different types of therapy during the course of their illness, but despite these treatment efforts none showed clinical improvement. The study was approved by the medical ethical committee of our hospital and took place between February 1996 and September 1997. All patients gave informed consent. No immunosuppressive drugs were administered within the 3 months prior to the study. IFN-β1a (Rebif®; Serono Benelux) was self-administered at a dosage of 6 million IU, three times a week for half a year and then, if clinical improvement was found (defined as at least one point improvement on the modified Rankin), the treatment was continued for an additional period of six months. To minimise the chance of adverse events a lower dose of 1.2 MIU Rebif was administered during the first week and 3.0 MIU during the second week. Thereafter the full dosage was given. Acetaminophen (500-1000 mg/day) was administered prophylactically during the first 6 weeks of treatment to ameliorate known constitutional symptoms of IFN-β1a.

Clinical assessment

Primary outcome was assessed at the disability level using the Nine-hole peg test, the Ten-metre walking test, and the modified Rankin scale (9-11). All patients received training in fulfilling the Nine-hole peg test before the start of the study to exclude any training effect. Secondary outcome was measured at the impairment level using the MRC sumscore, which was slightly modified (12). The following muscle pairs were examined: upper arm abductors, elbow flexors, wrist extensors, interosseus muscles, hip flexors, knee extensors, foot plantar flexors and foot dorsal flexors (score range: 0-80). All tests were assessed under pre-defined standard conditions. The scales were applied at entry and once a week in weeks 2, 4, 6, 8, 10, 12, 16, 21, 26 in all patients, three months after stopping IFN-β1a in two

patients and in weeks 32, 42, and 52 in the other two patients. Two investigators (IM/PD) did the follow-up assessments, each examining two patients. All measurements were compared with the baseline findings for each patient. Adverse events were recorded.

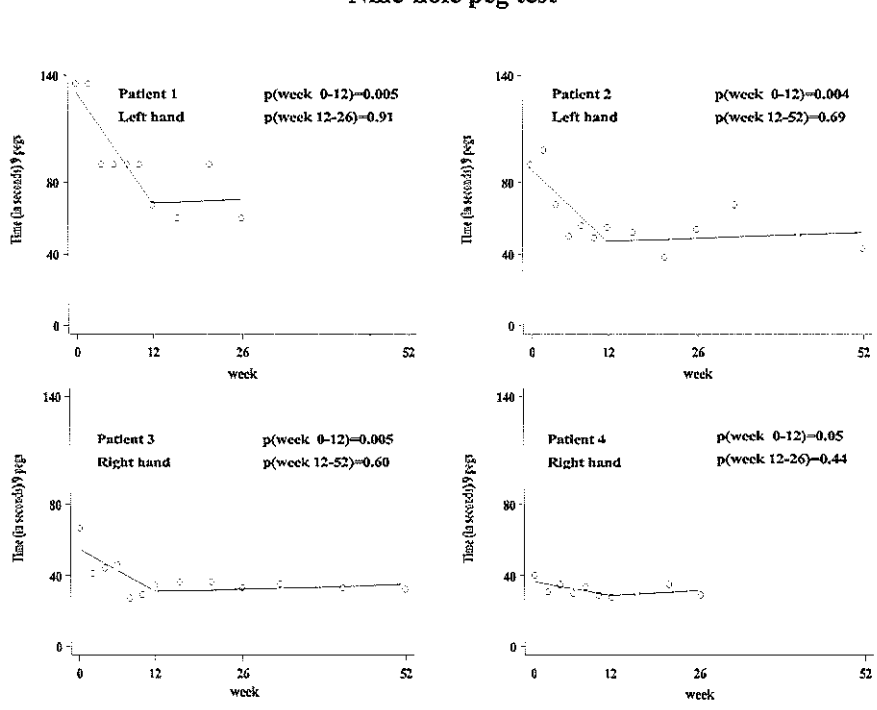
Additional investigations
Routine physical examination and laboratory studies, including enzyme linked immunosorbent assay (Elisa) tests for antibodies against the ganglioside GM1 were performed within two weeks before the start of the study and subsequently five times during treatment period (13). Electromyography was performed under standard conditions using supramaximal stimulation by the same examiner (JM) within two weeks before day 1, and consecutively 3 to 5 times during therapy time. Nerve conduction velocities and compound muscle action potential (CMAPs) were examined in eight motor nerves (four of the upper and four of the lower limbs). The examination always included the affected nerve(s) resulting in impairment.

Statistical analysis
Conventional linear and linear spline (piecewise method) regression analyses were used to evaluate the obtained serial data for the Nine-hole peg and the Ten-metre walking test (14). The knots of the linear spline functions were taken at week 12 of treatment, based on the clinical picture that was observed. This will be further discussed. All analyses were performed using Stata 5.0 for Windows 95 (Stata Statistical Software: Release 5.0. 702 University Drive East, College station, TX: Stata Corporation 1997). A p-value ≤ 0.05 was considered to be significant.

Results

After approximately 2 to 4 weeks of treatment the patients reported some improvement in walking and their daily manual skills. A maximum improvement was reached around 3 months of therapy, with hereafter a stabilisation or only minimal clinical improvement. None of the patients experienced deterioration during treatment.
Manual skills such as washing and brushing hair, dressing upper part of the body, doing/undoing buttons and zips, and opening a jar or a bottle were more easily accomplished. The time needed to fulfil the Nine-hole peg test by the most affected hand of each patient was significantly reduced in all patients during the first 12 weeks of treatment (figure 1). Improvement was also seen in the other less affected hands in the first 12 weeks, but this was significant only in patient number 3 ($p = 0.003$).
Improvement in ambulation consisted of a more easily walking pattern in all patients. All patients claimed to need less assistance of another person and used their aid(s) less then they were used to. An ability to walk for a longer time was also experienced by two patients. The time required to walk 10 meters was significantly reduced in three patients in the first 3 months (figure 2). Although ambulation improvement in patient number 3 was not significant, her endurance improved considerably within the first 3 months of treatment as she could walk for more than two hours in the woods. The Rankin score also changed notably in this patient, from 3 to 1 around three months of therapy. Although the other 3

Figure 1



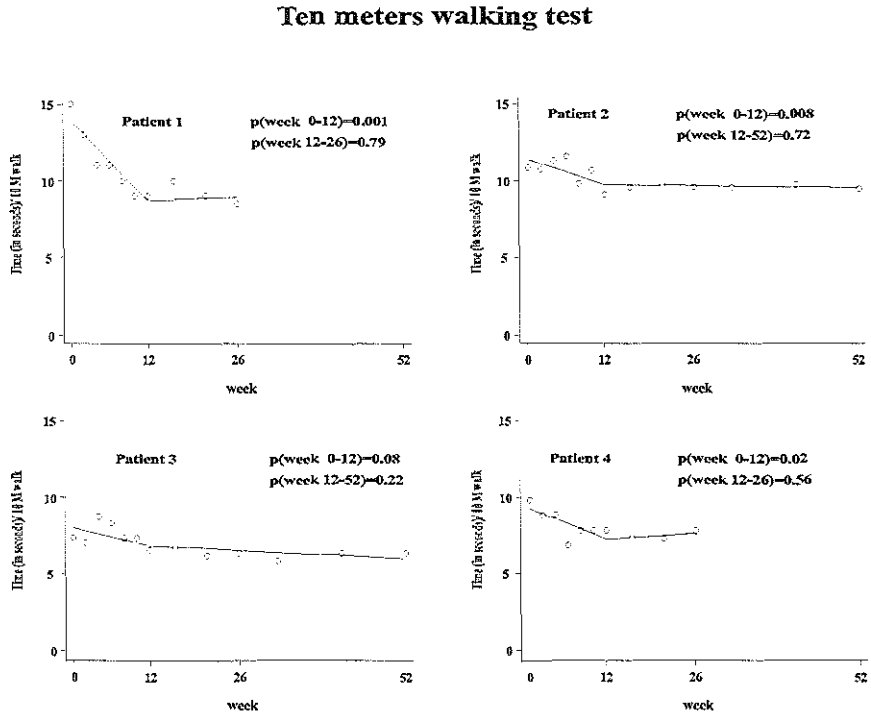
Legends to Figure 1
Most affected hand of each patient. Analysis performed using linear spline regression methods with the knots of the linear spline functions taken at 12 weeks.

patients showed improvement at the level of impairment and disability, the Rankin score only improved in patient number 2 (from 4→3). The Rankin score of the other two patients remained 3. Based on these results, we decided to discontinue IFN-β1a in the patients 1 and 4 with unchanged Rankin score after 6 months. Patient number 4 remained stable at all levels of measuring outcome during the next three months. Patient number 1 experienced slight deterioration in strength, dexterity, and mobility, but his Rankin score remained unchanged. The IFN-β1a treatment was continued for another 6 months in two patients (number 2 and 3).

Conventional linear regression analysis demonstrated significant improvement in muscle strength in all patients during the course of treatment ($p \leq 0.001$ for patients 1 to 3; $p = 0.04$ for patient 4). The MRC sumscore increased from 40 to 53, 53 to 60, 69 to 73 and 49 to 57, respectively in patients 1 to 4.

All patients had motor conduction blocks (MCBs), outside usual nerve compression sites in various nerves, ranging from 35% to 94%. Only the patient with chronic inflammatory demyelinating polyneuropathy had a marked partial decrease of the motor conduction blocks in the right ulnar (82%→37%), left ulnar (70%→24%), and left median nerve (94%→53%).

Figure 2



Legends to Figure 2
Analysis performed using linear spline regression methods with the knots of the linear spline functions taken at 12 weeks.

Motor nerve conduction velocities did not improve. Anti-GM1 antibody titres did not change. The recorded side effects of IFN-β1a were flu-like symptoms, fever, sweating, and erythema at the injection sites. These disappeared gradually within 2 months. The drug was well tolerated. Physical examination and routine blood and urine analysis remained normal.

Discussion

In this open prospective study, treatment with IFN-β1a (Rebif) induced clinical improvement in the first 3 months of therapy in all 4 patients with severe chronic motor neuropathies not improving after conventional therapies. All patients remained stable during the follow-up treatment period. However, improvement on the modified Rankin was only observed in two patients. A possible explanation for this finding is that the grading definitions of the modified Rankin are very broad classifications of disability and therefore rather insensitive to detect improvement as observed on the other scales used. A similar finding was noted in a recent publication studying the effect of IVIg in chronic inflammatory demyelinating polyneuropathy (15). Patient 3 seemed to be less profoundly affected than the

other patients. She especially had less severe amyotrophy, which probably explains her better score on the modified Rankin scale. The response to treatment, therefore, seems to be correlated with the degree of being affected and the severity of amyotrophy.

Improvement after IFN-β1a was also recently observed by Choudhary *et al.* in a patient with an 8-year history of a relapsing and remitting sensory-motor chronic inflammatory demyelinating polyneuropathy, not responding to various conventional treatments (16). This patient received 3 MIU of IFN-β1a, three times a week. Improvement began 2 weeks after administration and, as in our patients a maximum was reached after 12 weeks. Other reports have also shown a possible therapeutic effect of this group of regulatory cytokines in chronic immune-mediated neuropathies (17-19). The present study provides some support for the effectiveness of IFN-β1a in patients with chronic immune-mediated neuropathies, particularly in patients with multifocal motor neuropathy.

A poor correlation between clinical improvement and neurophysiological data, as seen in this study, has also been reported by others (5,6,20). One of the possible explanations for this discrepancy is fluctuation in temporal dispersion, which may result in alterations of the form of the CMAPs (20). Another possible cause is that restored conduction blocks located at the most proximal nerve segments may not be detected by neurophysiological studies.

The pathophysiologic mechanism of action of IFN-β in chronic immune-mediated neuropathies is not known. Presently, the knowledge regarding its immunoloigcal effects is mainly derived from studies on multiple sclerosis (21,22). IFN-β may counteract the effects of IFN-gamma such as downregulation of major histocompatibility (MHC) class II antigen expression on neuroendothelial cells (21,22). This may be of importance since upregulation of MHC class II molecules on endoneurial cells has been demonstrated in chronic inflammatory demyelinating polyneuropathy (23). Other immunomodulating effects of IFN-β that may be significant in multifocal motor neuropathy and chronic inflammatory demyelinating polyneuropathy include enhancement of T-suppressor cell function, reduction of T-cell activation and the production of IFN-gamma, downregulation of the production of certain cytokines such as tumour necrosis factor alpha, and induction of the production and secretion of interleukin-4 and interleukin-10 (21,22,24).

No severe adverse events were recorded and fortunately none of the patients deteriorated during the administration of IFN-β1a. This suggests that IFN-β1a presumably can be prescribed safely in patients with a chronic motor neuropathy. It is suggested, if patients respond to treatment with IFN-β1a, that improvement generally starts after 2-4 weeks. If there is no improvement after about 3 months, therapy should be discontinued.

In conclusion, our findings indicate that severely affected patients with chronic immune-mediated motor neuropathies not responding to conventional treatments may show improvement when treated with IFN-ß1a. A controlled trial is required to confirm these findings.

References

1. Pestronk A, Cornblath DR, Ilyas AA, *et al.* A treatable multifocal motor neuropathy with antibodies to GM1 ganglioside. Ann Neurol 1988;24:73-78.

2. Kornberg AJ and Pestronk A. Chronic motor neuropathies: Diagnosis, Therapy, and Pathogenesis. Ann Neurol 1995;37(suppl 1):S43-S50.

3. Oh SJ, Claussen GC, Odabasi Z, *et al.* Multifocal demyelinating motor neuropathy: Pathologic evidence of 'inflammatory demyelinating polyradiculoneuropathy'. Neurology 1995;45:1828-1832.

4. Krendel DA. Biopsy findings link multifocal motor neuropathy to chronic inflammatory demyelinating polyneuropathy. Ann Neurol 1996;40:948-949.

5. Chaudry V, Corse AM, Cornblath DR, *et al.* Multifocal motor neuropathy: Response to Human Immune Globulin. Ann Neurol 1993;33:237-242.

6. van den Berg LH, Kerkhoff H, Oey PL, *et al.* Treatment of multifocal motor neuropathy with high dose intravenous immunoglobulins: a double blind placebo controlled study. J Neurol Neurosurg Psychiatry 1995;59:248-252.

7. Feldman EL, Bromberg MB, Albers JW, *et al.* Immunosuppressive treatment in multifocal motor neuropathy. Ann Neurol 1991;30:397-401.

8. Report from an adhoc subcommittee of the American Academy of Neurology AIDS task force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

9. Mathiowetz V, Weber K, Kashman N, *et al.* Adult norms for the nine-hole peg test of finger dexterity. Occup Ther J Res 1985;5:24-38.

10. Collen FM, Wade DT, Bradshaw CM. Mobility after stroke: reliability of measures of impairment and disability. Int Disab Studies 1990;12:6-9.

11. van Swieten JC, Koudstaal PJ, Visser MC, *et al.* Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988;19:604-607.

12. Kleyweg RP, van der Meché FGA, Schmitz PIM. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré Syndrome. Muscle and Nerve 1991;14:1103-1109.

13. Jacobs BC, van Doorn PA, Schmitz PIM, *et al.* Campylobacter jejuni infections and anti-GM1 antibodies in Guillain-Barré Syndrome. Ann Neurol 1996;40:181-187.

14. Green WH. Econometric Analysis. Macmillan, New York. 2nd ed., 1993.

15. Thompson N, Choudhary P, Hughes RAC, *et al.* A novel trial design to study the effect of intravenous immunoglobulin in chronic inflammatory demyelinating polyradiculoneuropathy. J Neurol 1996;243:280-285.

16. Choudhary PP, Thompson N, Hughes RAC. Improvement following interferon beta in chronic inflammatory demyelinating polyradiculoneuropathy. J Neurol 1995;242:252-253.

17. Sabatelli M, Mignogna T, Lippi G, *et al.* Interferon-alpha may benefit steroid unresponsive chronic inflammatory demyelinating polyneuropathies. J Neurol Neurosurg Psychiatry 1995;58:638-639 (letter).

18. Mariette X, Chastang C, Clavelou P, *et al.* A randomised clinical trial comparing interferon-α and intravenous immunoglobulin in polyneuropathy associated with monoclonal IgM. J Neurol Neurosurg Psychiatry 1997;63:28-34.

19. Gorson KC, Ropper AH, Clark BD, *et al.* Treatment of chronic inflammatory demyelinating polyneuropathy with interferon-α 2a. Neurology 1998;50:84-87.

20. van Doorn PA, Brand A, Strenger PFW, *et al.* High-dose intravenous immunoglobulin treatment in chronic inflammatory demyelinating polyneuropathy: a double blind, placebo controlled, crossover study. Neurology 1990;40:209-212.

21. Weinstock-Guttman B, Ransohoff RM, Kinkel P, *et al.* The Interferons: Biological effects, mechanisms of action, and use in multiple sclerosis. Ann Neurol 1995;37:7-15.

22. Noronha A, Toscas A, Jensen MA. Interferon β decreases T cell activation and interferon γ production in multiple sclerosis. J Neuroimmunol 1993;46:145-154.

23. Pollard JD, McCombe PA, Baverstock J, *et al.* Class II antigen expression and T lymphocyte subsets in chronic inflammatory demyelinating polyneuropathy . J Neuroimmunol 1986;13:123-134.

24. Lugaresi A, De Luca G, Iarlori C, *et al.* Interferon beta normalises suppressor cell function in dysimmune neuropathies [abstract]. J Neurol 1997;244:S66.

Chapter 8

The development and clinimetric evaluation of a new handicap scale in immune-mediated polyneuropathies

Merkies ISJ[1], Schmitz PIM[2], van der Meché FGA[1], Samijn JPA[1], van Doorn PA[1], for the Inflammatory Neuropathy Cause And Treatment (INCAT) group[3]

From the Department of [1]Neurology, [2]Department of Statistics - Daniel den Hoed Cancer Centre, Erasmus University Medical Centre Rotterdam, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands. [3]See Chapter 2 for INCAT-members.

Submitted for publication

Abstract

Objectives: Development and clinimetric evaluation of a new handicap scale in immune-mediated polyneuropathies.
Methods: The Rotterdam nine items handicap scale (RIHS9) was constructed based on literature evaluation of potential items using the World Health Organisation Handicap scale as a framework, suggestions and judgement by patients who suffered from Guillain-Barré syndrome (GBS), chronic inflammatory demyelinating polyneuropathy (CIDP), or a gammopathy related polyneuropathy (MGUSP), and clinical experience by a panel of neurologists. Subsequently, the RIHS9 and two other measures (Rankin scale, Hughes' functional grading scale [f-score]) were assessed in 113 clinically stable patients (GBS 83, CIDP 22, MGUSP 8). The RIHS9 and Rankin scale were also utilised longitudinally in twenty patients with recently diagnosed GBS (n=7) or CIDP (n=13) and changing clinical conditions. The validity of the RIHS9 was examined in patients with various degrees of disease severity (f-score subgroups) (discriminant ability) and by correlation studies with the Rankin scale. The reliability of the RIHS9 was determined by calculating the inter-/intra-observer agreements and its responsiveness was analysed using the standardised-responsive-mean (SRM) technique.
Results: Good discriminatory abilities were obtained for the RIHS9. Patients with a worse clinical state (higher f-score values) had significantly more handicap disturbances (lower RIHS9 values). Good correlation was found between the RIHS9 and Rankin scale (stable group: Spearman's Rank test: r=-0.76 to -0.78; longitudinal group: Intraclass correlation coefficient: R=0.83; $p<0.0001$ for all associations). The RIHS9 demonstrated good reliability (R=0.89-0.98; $p<0.0001$) and high responsiveness values (SRM>0.8). In contrast with the Rankin scale, the RIHS9 provided not only information regarding mobility, but also highlighted the dimensions physical independence, occupation, and social integration.
Conclusion: The Rotterdam nine items handicap scale is demonstrated to be simple, valid, reliable, and responsive in patients with immune-mediated polyneuropathies. Its use is therefore suggested for evaluating outcome at the handicap level in these conditions.

Introduction

Defining and evaluating impairments and disabilities can be useful for diagnosis, implementation of a specific therapy or rehabilitation intervention, and understanding how illness has its impact on patients' daily lives (1,2). It is therefore not surprising that patients with Guillain-Barré syndrome (GBS), chronic inflammatory demyelinating polyneuropathy (CIDP) or a polyneuropathy associated with a monoclonal gammopathy of undetermined significance (MGUSP) have been traditionally investigated at the impairment and disability levels of outcome (2,3). However, these disorders tend to have a long-term adverse impact on patients' lives and it was therefore argued that more emphasis should also be directed towards the evaluation of patients' participation in social activities (4-7). Assessing outcome according to the World Health Organisation (WHO) at the level of 'handicap' (soon to be renamed 'participation'), as the ultimate goal of societal integration, would eventually provide a more comprehensive view on the functionality of patients, since handicap incorporates mobility, physical dependence, economic self-sufficiency, occupation, social participation, and social norms (2,7). For people with impairments or disabilities, the successful return to activities like travelling, fulfilling a job, performing a hobby, and interacting with others in daily life could be seen as the biggest goal in the process of recuperation. The WHO defines handicap as a disadvantage for a given individual, resulting from impairment or disability that limits or prevents the fulfilment of a role that is normal (depending on age, sex and social and cultural factors) for the individual (2).
Surprisingly, to date only one study in patients with immune-mediated polyneuropathies examined outcome at the handicap level (4). Several other papers examining these and other forms of polyneuropathy have claimed the assessment of handicap using the (modified) Rankin scale (8-12). However, the Rankin 'handicap' scale is a global functional index with a strong emphasis on physical disability, in particular mobility, and is therefore not suitable for assessing handicap, as defined by the WHO (2,13).
By reviewing the English literature, only a few instruments have been reported that purely assess handicap (7,14-17). The advantages and in particular their shortcomings have been recently highlighted (7). Also, none of these outcome measures have been evaluated in terms of their psychometric soundness in patients with immune-mediated polyneuropathies (18,19). The aim of our study was to construct a new handicap scale and to evaluate this measure for its time to be completed, its validity, reliability, and responsiveness in patients with immune-mediated polyneuropathies (18,19).

Methods

Phase I
International criteria for scale development.
International criteria for the development of outcome measures have been postulated and were recently thoroughly described by Streiner and Norman (18-21). These criteria are:
- Devising the items: literature search and critical review through key informant interviews of potential items that might be relevant for the construction of an outcome measure, in this case items that reflect handicap activities.

- The constructed scale should be easily applicable, requiring a minimum time to be completed (18,19);
- Also, it should be valid, reliable, and responsive to relevant changes over time (18,19).

By reviewing the literature, we recruited items that might plausibly be included in the new handicap scale (2,14-16). The ICIDH taxonomy was used as a foundation (2). In particular, the handicap dimensions "mobility", "physical independence", "occupation", and "social integration" were considered, thereby avoiding overlap with impairment and disability areas (2). Various items related to these dimensions were chosen taking into account the clinical spectrum of sensory-motor immune-mediated polyneuropathies (3). Efforts were made to describe most items in a simplistic, non-ambiguous, way. Subsequently, an open telephone interview was performed with twelve patients (GBS 9, CIDP 3) asking questions related to handicap activities and checking whether the selected items were relevant to them. Also, these items were mailed to a group of thirty-eight (GBS 25, CIDP 9, MGUSP 4) patients that judged each selected item separately on its significance in measuring part of their social activities. If needed, items were added to or excluded from the selected item-list. All patients fulfilled these requests and the obtained information was examined and if necessary incorporated to form the first draft of the Rotterdam Handicap scale.

In addition, this draft was thoroughly evaluated by a group of 13 senior neurologists, all members of the Inflammatory Neuropathy Cause And Treatment (INCAT) group, a collaborating force of European neurologists with special interest in neuro-immunological illnesses. The clinical applicability and relevance of all selected items were analysed, based on their personal experience and clinical view. Changes were made according to their suggestions. To strive for clarity, the experts suggested a distinction between tasks or activities that were fulfilled indoors or outdoors, because it was believed that these might require different efforts from the patient. For example, leisure activities indoors (such as "reading a newspaper, using telephone") would most probably require different efforts from the patient to accomplish compared with leisure activities outdoors (such as "going to meetings, theatre, concert, etcetera"). Also, guidelines for specific patients' related circumstances were provided and eventually the final form of the Rotterdam 9 Items scale (RIHS9) was presented for psychometric evaluation in immune-mediated polyneuropathies (see Appendix III, page 213-216).

Phase II

Patients

One hundred and thirteen patients (83 GBS, 22 CIDP, 8 MGUSP) with a stable clinical condition were recruited from the Rotterdam immune-mediated polyneuropathy databank and the Dutch GBS study group (stable group). Patients with GBS, CIDP or MGUSP were recruited, since it was argued that these disorders represent parts of a continuum regarding their neuromuscular dysfunction pattern (3). The selected patients still had residual symptoms or signs due to their illness, representing a broad range of disability. Nine CIDP patients required interval treatment ranging from weeks to months, with intravenous immunoglobulin (IVIg). With this therapy their clinical condition has been stable for more than 6 months. Six patients with MGUSP (three with IgG, two with IgM, and one patient with IgG+IgM) had an associated demyelinating polyneuropathy with minor concurrent axonal damage in three. An axonal polyneuropathy was diagnosed in the remaining two patients with MGUSP (one IgA and one IgG type).

Twenty consecutive patients with recently diagnosed sensory-motor GBS (n = 7) or CIDP (n = 13) and changing clinical conditions were enrolled to investigate the responsiveness of the RIHS9 (longitudinal group). All GBS and CIDP patients met the international criteria for their illness (22,23). The diagnosis MGUSP was established after excluding all possible causes for the gammopathy and polyneuropathy (24).

Assessment tools/scales description

The *Rotterdam 9 Items handicap scale* (*RIHS9*) comprises 9 inquiries with answers ranging from 1 ("unable to fulfil tasks/activities") to 4 ("complete fulfilment of tasks/activities"). Since not all items are necessarily applicable to all patients, an answer "0" ("not applicable") was added to the nine inquiries. To avoid the formation of various subgroups of patients with different amounts of applicable items (for example: a subgroup with 9 applicable items [raw score-range: 9-36], another with 8 [raw score-range: 8-32], etcetera), all initial raw scores were multiplied by 9/(9-number of not applicable items). Hence, the final RIHS9 score-range was independent of the amount of applicable items and extended from 9 ("unable to fulfil any applicable task or activity") to 36 ("able to fulfil all applicable tasks or activities").

The Modified *Hughes' functional grading scale* (*f-score*) assesses the functional ability of patients with a strong emphasis on mobility (25). The f-score of the patients included in this study ranged from 0 to 5. An f-score of 0 indicates normal health (no symptoms or signs); 1, denotes having minor neurological symptoms or signs and able to run; 2, able to walk at least 10 meters, but unable to run; 3, able to walk 10 meters with a walker or support; 4, bedridden or chair bound (unable to walk 10 meters with a walker or support); 5, requiring artificial ventilation for at least part of the day (25). Good psychometric requirements were fulfilled for this scale (25).

The *Rankin scale* has been primarily used in patients with stroke (8). The grades of this scale range from 0 to 5. A Rankin score 0 indicates no symptoms at all; 1, no significant disability despite symptoms: able to carry out all usual duties and activities; 2, slight disability: unable to carry out all previous activities but able to look after own affairs without assistance; 3, moderate disability: requiring some help, but able to walk without assistance; 4, moderately severe disability: unable to walk without assistance or unable to attend to own bodily needs without assistance; and 5, severe disability: bedridden, incontinent, and requiring constant nursing care and attention (8). No formal psychometric evaluation has been performed in patients with immune-mediated polyneuropathies.

Psychometric test procedures

General aspects. All participants gave informed consent before the study. All measures were obtained in a quiet and comfortably warm room at our outpatient clinic. The assessments were performed in a random order. The time to complete the RIHS9 was recorded at each assessment (in seconds). The study took place between December 1998 and January 2000 and was performed on behalf of the INCAT group".

Validity and Reliability. Various aspects of validity were evaluated for the RIHS9 (18,19). "Validity by assumption" (i.e., face and content validity) was obtained for this scale, by reviewing and judging all aspects of the RIHS9 by the INCAT expert panel of neurologists (18,19). The discriminant validity of the RIHS9 was investigated by examining various subgroups of stable patients with different degrees of clinical state. The f-score was used to

form the various subgroups (25). In particular, we examined whether there would be a difference in handicap, as assessed by the RIHS9, when comparing patients who could walk independently (f-score ≤ 2) versus those who required help or were bed or chair bound (f-score > 2). Also, the correlation between the RIHS9 and Rankin scale was evaluated in both stable and longitudinal groups of patients.

For the assessment of reliability in the stable group of 113 patients, two neurologists and six experienced residents in neurology formed 28 different couples. Preceding the study, all investigators received instructions in assessing the RIHS9. Twenty-seven ("variable") couples investigated a total of 68 patients (2-3 patients for each couple). The remaining 45 stable patients were examined by the "experienced" couple (IM+JS). The latter couple was formed to examine the effect of training (and thus a possible increase in reliability) when using the RIHS9 often.

The patients were examined on two different occasions at our outpatient clinic. During the first visit the two members of an appointed pair performed their scores independently and consecutively (usually within 2 hours) (inter-observer measures). Within 2-4 weeks, the patient returned for a second visit and only one investigator of the earlier assigned pair examined the patient again (intra-observer values) without having access to previous results. The assessments sequence at entry and the examination at the second visit were equally distributed among the members of an assigned couple. Eventually, each member of a couple examined approximately the same number of patients. With exception of the f-score, the RIHS9 and Rankin scale were assessed at each visit in all patients. For the validity studies, only the recruited scales' values at one visit were used.

Responsiveness. Twenty consecutive patients were longitudinally examined by the same clinician (IM). The RIHS9 and Rankin scale were assessed at study entry and at the weeks 2, 4, 8, 12, 16, 21, 26, 32, 40, and 52 of follow-up with additional clinical investigations if necessary.

Statistics

Validity and Reliability. In the stable group, the discriminant validity of the RIHS9 was investigated by quantile regression analyses on the f-score patients' subgroups (f-score = 1 to f-score = 4) (26). The median RIHS9 scores in these subgroups were compared, using design variables (0 to 1) representing each subgroup. The design variables were defined in an appropriate manner, making it possible to test (by the likelihood ratio test) the hypothesis that the median RIHS9 values were different for all f-score subgroups and decreased with increasing f-score values. Box plots, with the upper and lower adjacent values defined according to Tukey, were applied to visualise these observations (also see legend to Figure for explanation of the various components of a box plot) (27). Also, differences in median RIHS9 values and at the single-item level of this scale were investigated in patients who could walk independently (f-score ≤ 2) versus those being dependent or unable to walk (f-score > 2) (Mann-Whitney U test).

In the stable group, the correlation between the RIHS9 and Rankin scale was examined using Spearman's Rank correlation test. Also, random effects linear regression analyses were performed of the RIHS9 on the Rankin scale in the longitudinal group, taking into account the association of the data caused by the longitudinal structure. The latter was achieved using the program "xtreg" in STATA 6.0 (StataCorp. 1997. Stata Statistical Software: Release 6.0. College Station, TX), which is based upon a cross-sectional time-

series regression model as described by Dwyer and Feinleib (28). The inter- and intra-rater reliability of the RIHS9 was quantified by estimation of the intraclass correlation coefficient using a one-way random effects analysis-of-variance model (Anova) for the two investigator ("experienced" and "variable") groups.

Responsiveness. Responsiveness was investigated by calculating the standardised response mean (SRM) scores for the RIHS9 at various arbitrarily chosen occasions during follow-up (weeks 12, 26, 40, 52) (29). SRM is equal to the mean change in score divided by the standard deviation of the change in score (29). A value between 0.5 and 0.8 is considered moderate, and 0.8 or greater as high responsiveness (29,30). All analyses were performed using Stata 6.0 for Windows 95. A value of $p < 0.05$ was considered statistically significant.

Table 1

Basic characteristics of patients with immune-mediated polyneuropathies

Stable group of patients (n = 113; GBS 83, CIDP 22, MGUSP 8)	
Median age at start of the study (range) [years]	56 (14 – 84)
Median Rotterdam 9 Items Handicap scale (score range: 9 – 36)	
At entry	31.5 (14 – 36)
Second visit	31.5 (13 – 36)
Third visit	31.5 (16 – 36)
Median Rankin score at entry (score range: 0 – 5)	2 (0 – 4)
Hughes' functional grading scale at entry (score range: 1 - 4)	
1	51 (45%)
2	41 (36%)
3	14 (12%)
4	7 (6%)
Longitudinal group of patients (n = 20; GBS 7, CIDP 13)	
Median age at start of the study (range) [years]	54 (15 – 70)

Legend to Table 1
GBS = Guillain-Barré syndrome; CIDP = chronic inflammatory demyelinating polyneuropathy; MGUSP = polyneuropathy associated with a monoclonal gammopathy of undetermined significance. *INCAT* = inflammatory neuropathy cause and treatment group.

Results

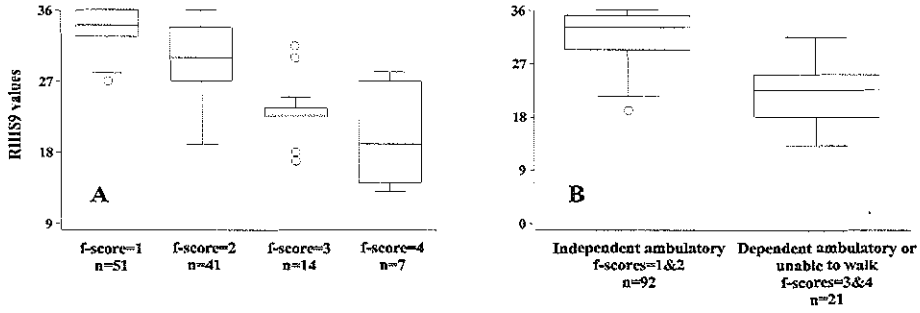
General aspects. The whole effort of literature review, recruitment of potential items, interviews and judgements by the patients of the selected items, and evaluation by an expect panel of neurologists, was accomplished in approximately six months with the Rotterdam 9 Items Handicap Scale (RIHS9) as the final outcome (see Appendix III, page 213-216). Subsequently, the RIHS9 was psychometrically evaluated in the selected patients in the current study. A median time of 3.5 (range: 0.5 to 11.5) minutes was needed to complete the RIHS9.

The basic characteristics of the patients selected for the psychometric study of the RIHS9 are presented in Table 1. The stable group of patients (54 females and 59 males) had a median duration of symptoms till onset of the study of 5.1 years. Seven of these patients were bed/chair bound and fourteen required assistance or a device to walk short distances. The remaining 92 patients could walk independently. The corresponding values and ranges for the scales in these patients are presented in Table 1.

Validity, inter- and intra-observer Reliability. By the likelihood ratio test, it was demonstrated that the median RIHS9 values were different between all subgroups (Figure A). This test also revealed that the subgroup f-score=1 had a higher median RIHS9 value compared with the other subgroups (Figure A; f-score=1: 34; f=score=2: 30; f-score=3: 23; f-score=4: 19; quantile regression analyses combined with likelihood ratio test: $p < 0.0001$ for all comparisons). The median RIHS9 value in the independent ambulatory subgroup of patients (f-score ≤ 2) was significantly higher than the median value for dependent subgroup (f-score 3 & 4) (Figure B; independent ambulatory: median RIHS9 value 33; dependent: 22.5; Mann Whitney U test: $p < 0.0001$). This difference was also reflected in all item-level

Figure

Rotterdam 9 Items handicap scale scores (RIHS9) in stable patients with immune-mediated polyneuropathies with various levels of clinical state



Legend to Figure

Box plots demonstrating the obtained Rotterdam 9 items handicap scale (RIHS9) scores in the various subgroups of patients with different degrees of disease severity (A: based on the f-scores1 to 4; B: based on independent ambulatory [f-score ≤ 2] versus dependent ambulatory/unable to walk subgroups [f-score > 2]). As noted in Figure A: The median RIHS9 values decreased significantly with increasing f-score (worse clinical state) (Likelihood Ratio-test combined with quantile regression analysis: $p < 0.0001$ for all comparisons). The independent ambulatory subgroup of patients demonstrated a higher median RIHS9 score compared with the dependent ambulatory/not able to walk subgroup (Mann Whitney U test: $p < 0.0001$).

comparisons (Table 2). At all item levels, the dependent ambulatory subgroup of stable patients (f-score > 2) had significantly lower values (indicating more handicap) compared with the independent ambulatory patients (Mann-Whitney U test: $p \leq 0.04$ for all comparisons).

In the stable group, 24% of the patients returned to their previous full-time job or study (item "being able to work or to study"). In the longitudinally examined patients, 20% of the patients were able at entry to fulfil their previous work or study as before. At one-year follow-up, this number increased to 60%.

In the stable group of patients, a significant correlation was seen between the RIHS9 and the Rankin scale (Spearman's Rank correlation test, for the "experienced" couple: $r = -0.78$; for the "variable" couples: $r = -0.76$; $p < 0.0001$ for both correlations). A good correlation was

Table 2

Single-item comparisons in a group of stable patients with immune-mediated polyneuropathies and various degrees of clinical state

Rotterdam 9 Items handicap scale	Independent ambulatory subgroup (f-score ≤ 2) (n = 92) *	Needing support to walk/unable to walk subgroup (f-score > 2) (n = 21) *
Items	Percentage of patients with maximum score (%)	Percentage of patients with maximum score (%)
Mobility indoors	92/92 (100)	20/21 (95)
Mobility outdoors	81/92 (88)	4/20 (20)
Kitchen tasks	68/89 (76)	3/16 (19)
Domestic tasks indoors	59/86 (69)	0/12 (0)
Domestic tasks outdoors	50/90 (56)	1/18 (6)
Leisure activities indoors	70/92 (76)	12/21 (57)
Leisure activities outdoors	39/90 (43)	0/19 (0)
Able to drive a car–go by bus–ride a bicycle	68/92 (74)	0/20 (0)
Able to work or study	17/59 (29)	0/11 (0)

Legend to Table 2
Maximum score = able to fulfil all tasks or activities related to the question asked. For the description of the items: see appendix III, pages 213-216; * implies that not all items were necessarily applicable to all patients. Therefore, the amount of patients per each item could vary.

also obtained between these scales in the longitudinally examined patients ("xtreg": R = 0.83; $p < 0.0001$). Good reliability values were demonstrated for the RIHS9 (Anova, "experienced" couple: inter-observer: R=0.89; intra-observer: R = 0.93; "variable" couples, inter-observer: R = 0.94; intra-observer: R = 0.98; $p < 0.0001$ for all associations).

Responsiveness. Eight females and twelve males were examined longitudinally (Table 1). At study entry, four patients were bed bound, one requiring artificial ventilation, and nine patients were unable to walk independently. All patients experienced deficit in daily functional activities. Two hundred and one visits were completed during a follow-up period of 40 to 58 (median: 52) weeks. Nineteen patients completed a one-year follow-up. With the exception of one GBS patient who only experienced mild symptoms, all patients received initial treatment with IVIg (0.4 grams/kilogram body weight/day for 5 consecutive days). All but one patient with CIDP showed good functional improvement on IVIg during follow-up. The non-responder received a treatment course of oral prednisone, 100 milligrams/day for four consecutive weeks. This patient also improved with this therapy and prednisone was tapered down over five months period to 30 milligrams on alternate days. The GBS patients did not show any deterioration during follow-up and improved gradually over time. After initial improvement, all 12 IVIg responsive CIDP patients needed interval therapy with IVIg (1-2 days 0.4 grams/kg/day; intervals: 3 - 21 weeks) to maintain the achieved improvement. Eventually, all patients demonstrated during follow-up an improvement in functional abilities. This was reflected in an increase of the RIHS9 values during follow-up (median values: 33, 35, 35, 34.5, at weeks 12, 26, 40, and 52, respectively) compared with the median entry value of 20 (Mann Whitney U test: $p \leq 0.001$ for all comparisons). Good SRM scores were obtained for the RIHS9 in these patients (SRM values: 1.1, 1.4, 1.3, and 1.4 at weeks 12, 26, 40, and 52).

Influence of sex, age, duration of symptoms, and diagnosis on handicap. In the stable group of patients, sex and duration of symptoms did not have a significant impact on RIHS9 values

($p > 0.09$). However, the RIHS9 was significantly related to age with a gradual decrease with increasing age (Spearman's Rank correlation: $r = -0.23$; $p = 0.02$). No differences in RIHS9 values were obtained when comparing patients with GBS versus those with a chronic course (CIDP/MGUSP).

Discussion

In the current study, the development and psychometric evaluation of a new handicap scale, the Rotterdam 9 items handicap scale (RIHS9), is presented for patients with immune-mediated polyneuropathies. Its development was performed under the prescribed international guidelines for scale construction using the ICIDH taxonomy as a foundation (18-21). The RIHS9 demonstrated itself to be easily applicable and only required a median time of 3.5 minutes for the evaluation of various handicap items in these conditions. Because the validity, reliability, and responsiveness of the RIHS9 were also demonstrated in these disorders, all psychometric essentials were fulfilled (19). The RIHS9 clearly distinguished between patients with various degrees of disease severity. Increasing f-score values (compatible with more disablement) in patients with immune-mediated polyneuropathies was accompanied by more handicap disturbances (lower RIHS9 values; see Figure A). Furthermore, the reproducibility of the RIHS9 turned out to be independent of the frequent use of this outcome measure, because the reliability values did not differ between the "experienced" and "variable" investigative couples. The RIHS9 provided, in contrast with the Rankin scale, not only information regarding mobility, but also highlighted the dimensions physical independence, occupation, and social integration.

To date, despite the daunting array of available clinical studies in patients with immune-mediated polyneuropathies and the growing aim of neurologists to ameliorate handicap on the long-term, only one study in these disorders have reported on handicap disadvantages (4). In this study, patients with Guillain-Barré syndrome were examined using the Environmental status scale (ESS) and the Handicap assessment scale (HAS) to evaluate outcome at the handicap level (4,31). However, using the ESS can be misleading, because this scale mixes disability and handicap measures and also demonstrated limited validity (31). Moreover, information regarding the psychometric requirements for the HAS have not been provided before its use in these conditions. To our knowledge, the current study is the first to assess handicap in patients with immune-mediated polyneuropathies using a scale that has been thoroughly psychometrically evaluated (19).

By reviewing the English literature, only three instruments have been published that purely assess handicap (14-16). None of these scales have been evaluated in immune-mediated polyneuropathies. Whiteneck and colleagues quantified handicap using their Craig handicap assessment and reporting technique (CHART) (14). The validity and reliability of this CHART were demonstrated in patients with spinal cord injury (14). However, specific responsiveness evaluation has not been reported. Moreover, the physical independence subscale of the CHART has been criticised for its assumption that all patients have access to attendant care, and therefore a patient who does not have help would be scored as independent (7,15). The community integration questionnaire (CIQ) was developed and psychometrically evaluated in patients with brain injury (15). However, like the CHART, the CIQ is a behavioural measure of handicap, since it only measures the time that a patient is

engaged in a particular activity and does not assess whether particular activities or interactions were successful accomplished (7). The London handicap scale assesses handicap as outlined by the WHO (2). Its validity and reliability have been demonstrated in the healthy elderly, patients with stroke and in inflammatory arthritis (16,17,32). The Effect size of this scale, a measure of responsiveness, was examined in patients with arthritis, but turned out to be relatively low (32). Moreover, the dimensions "looking after yourself" and "work and leisure" are constructed by a mixture of tasks and activities. In our view, these tasks and activities would probably require different efforts by patients to fulfil and might therefore lead to different degrees of handicap disturbances. It is therefore conceivable that the fulfilment of these enquiries would lead to some confusion. To strive for clarity and to group items that most probably would require equivalent efforts to fulfil, indoors and outdoors tasks and activities were defined in the RIHS9, hence providing a more simplistic view of patients' disadvantages at the handicap level.

In the current study, the RIHS9 revealed that only 24% of the stable patients resumed their previous employment or study. In the longitudinal group, this percentage shifted from 20% at entry to 60% at one-year of follow-up. Similar findings in alteration of earlier work were reported in patients with GBS (5,6,33). Age had an inverted correlation with handicap in the patients at study. This is in accordance with an earlier paper (34).

With respect to the aims of the current study, some methodological issues should be addressed. First, the obtained SRM scores for the RIHS9 only demonstrated within-group responsiveness. It is not clear whether substantial longitudinal discriminative responsiveness scores will be obtained for the RIHS9 when evaluating various groups of patients, e.g. in a trial setting comparing a placebo versus a treated group (35). Second, the WHO handicap dimensions "orientation" and "economic self-sufficiency" were not incorporated in the RIHS9 development (2). "Orientation" was not incorporated, because immune-mediated polyneuropathies, as peripheral nervous system disorders, in generally do not cause any alteration in this dimension. Moreover, the interviewed patients did not address at any time this entity as being important. "Economic self-sufficiency", the ability of a patient to sustain customary socio-economic activity and independence, was initially incorporated in the first draft of the RIHS9 (2). However, this item was excluded from the final version of this scale, because it turned out to be very difficult to measure. Furthermore, it was believed that this item had an overlap with the item "work/study".

In conclusion, the development, simplicity, validity, reliability, and responsiveness are demonstrated for the Rotterdam 9 items handicap scale (RIHS9) in patients with immune-mediated polyneuropathies. The RIHS9 adequately monitored the societal disadvantages in patients suffering from these disorders and seems therefore suitable for assessing handicap in these conditions.

References

1. Harwood RH, Jitapunkul S, Dickinson E, *et al.* Measuring handicap: motives, methods, and a model. Qual Health Care 1994;3:53-57.

2. World Health Organisation. International classification of impairments, disabilities, and handicaps. Geneva: WHO, 1980.

3. Hartung H-P, van der Meché FGA, Pollard JD. Guillain-Barré syndrome, CIDP and other chronic immune-mediated neuropathies. Curr Opin Neurol 1998;11:497-513.

4. Nicholas R, Playford ED, Thompson AJ. A retrospective analysis of outcome in severe Guillain-Barré syndrome following combined neurological and rehabilitation management. Disabil Rehabil 2000;22:451-455.

5. Bernsen RAJAM, de Jager AEJ, Schmitz PIM, *et al.* Residual physical outcome and daily living 3 to 6 years after Guillain-Barré syndrome. Neurology 1999;53:409-410.

6. de Jager AEJ, Minderhoud JM. Residual signs in severe Guillain-Barré syndrome: analysis of 57 patients. J Neurol Sci 1991;104:151-156.

7. Coulthard-Morris L, Burks JS, Herndon RM. Rehabilitation outcome measures. In: Herndon RM, editor. Handbook of neurologic rating scales. New York: Demos Vermande, 1997, 225-264.

8. van Swieten JC, Koudstaal PJ, Visser MC, *et al.* Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988;19:604-607.

9. Van Doorn PA, Vermeulen M, Brand A, *et al.* Intravenous immunoglobulin treatment in patients with chronic inflammatory demyelinating polyradiculoneuropathy. Clinical and laboratory characteristics associated with improvement. Arch Neurol 1991;48:217-220.

10. Hattori N, Ichimura M, Aoki S, *et al.* Clinicopathological features of chronic inflammatory demyelinating polyradiculoneuropathy in childhood. J Neurol Sci 1998;154:66-71.

11. Vermeulen M, van Doorn PA, Brand A, *et al.* Intravenous immunoglobulin treatment in patients with chronic inflammatory demyelinating polyneuropathy: a double blind, placebo controlled study. J Neurol Neurosurg Psychiatry 1993;56:36-39.

12. Teunissen LL, Eurelings M, Notermans NC, *et al.* Quality of life in patients with axonal polyneuropathy. J Neurol 2000;247:195-199.

13. de Haan R, Limburg M, Bossayt P, *et al.* The clinical meaning of Rankin 'handicap' grades after stroke. Stroke 1995;26:2027-2030.

14. Whiteneck GG, Charlifue SW, Gerhart KA, *et al.* Quantifying handicap: a new measure of long-term rehabilitation outcomes. Arch Phys Med Rehabil 1992;73:519-526.

15. Willer B, Rosethal M, Kreutzer JS, *et al.* Assessment of community integration following rehabilitation for traumatic brain injury. J Head Trauma Rehab 1993;8:75-85.

16. Harwood RH, Rogers A, Dickinson E, *et al.* Measuring handicap: the London Handicap Scale, a new outcome measure for chronic disease. Qual Health Care 1994;3:11-16.

17. Harwood RH, Gompertz SE. Handicap one year after a stroke: validity of a new scale. J Neurol Neurosurg Psychiatry 1994;57:825-829.

18. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. New York, Oxford university press, 1998.

19. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. J Neurol Neurosurg Psychiatry 1996;60:127-130.

20. Bombardier C, Tugwell P. A methodological framework to develop and select indices for clinical trials: statistical and judgmental approaches. J Rheumatol 1982;9:753-757.

21. Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chron Dis 1985;38:27-36.

22. Asbury AK and Cornblath DR. Assessment of current diagnostic criteria for Guillain-Barré Syndrome. Ann Neurol 1990;27(suppl):S21-S24.

23. Ad Hoc subcommittee of the American Academy of Neurology AIDS task force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

24. Miescher GC, Steck AJ. Paraproteinaemic neuropathies. Baillieres Clin Neurol 1996;5:219-232.

25. Kleyweg RP, van der Meché FGA, Schmitz PIM. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré syndrome. Muscle Nerve 1991;14:1103-1109.

26. Gould WW, Rogers WH. Quantile regression as an alternative to robust regression. 1994 proceedings of the statistical computing section. Alexandria, VA: American Statistical Association, 1994.

27. Tukey JW. Exploratory Data Analysis. Reading, MA: Addison-Wesley Publishing Company, 1977.

28. Dwyer J, Feinleib M. Introduction to statistical models for longitudinal observation. In: Dwyer J, Feinleib M, Lippert P, Hoffmeister H, editors. Statistical models for longitudinal studies of health. New York: Oxford University Press, 1992; 3-48.

29. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990;28:632-42.

30. Cohen J. *Statistical power analysis for the behavioural sciences*. New York: Academic Press; 1977.

31. Steward G, Kidd D, Thompson AJ. The assessment of handicap: an evaluation of the environmental status scale. Disabil Rehabil 1995;17:312-316.

32. Harwood RH, Carr AJ, Thompson PW, *et al.* Handicap in inflammatory arthritis. Br J Rheumatol 1996;35:891-897.

33. Lennon SM, Koblar S, Hughes RAC, *et al.* Reasons for persistent disability in Guillain-Barré syndrome. Clin Rehabil 1993;7:1-8.

34. Harwood RH, Prince M, Mann A, *et al.* Associations between diagnosis, impairments, disability, and handicap in a population of elderly people. Int J Epidemiol 1998;27:261-268.

35. Guyatt G, Walter S, Norman G. Measuring change over time: Assessing the usefulness of evaluative instruments. J Chron Dis 1987;40:171-178.

Chapter 9

Fatigue in immune-mediated polyneuropathies

Merkies ISJ[1], Schmitz PIM[2], Samijn JPA[1], van der Meché FGA[1], van Doorn PA[1], for the Inflammatory Neuropathy Cause And Treatment (INCAT) group[3]

From the Department of [1]Neurology, [2]Department of Statistics - Daniel den Hoed Cancer Centre, Erasmus University Medical Centre Rotterdam, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands. [3]See Chapter 2 for INCAT-members.

Neurology 1999;53:1648-1654; published with permission from the American Academy of Neurology

Abstract

Objectives: To determine the prevalence and severity of ongoing fatigue and to investigate the internal consistency, test-retest reliability, and validity of the Fatigue Severity Scale (FSS) in patients with immune-mediated polyneuropathies.
Methods: The FSS was assessed in 113 patients who either experienced Guillain-Barré syndrome in the past or currently have a stable chronic inflammatory demyelinating polyneuropathy or a polyneuropathy associated with a monoclonal gammopathy of undetermined significance, and in 113 age and sex matched healthy controls. Data on five additional scales (MRC sumscore, functional grading scale [f-score], *INCAT* sensory sumscore, a short form fatigue scale, medical outcome study 36-items health survey [SF-36]) were obtained in all patients. SF-36 was also assessed in 59 controls.
Results: 'Severe' fatigue (= FSS scores $\geq 95^{th}$ percentile values in controls) was present in 80% of the patients. Fatigue was not significantly related to general strength, sensory deficits, f-score, and duration of symptoms. 'Severe' fatigue was reported in 81 to 86% of patients with normal strength or sensation. Eighty percent of the patients (controls: 12%) reported their fatigue being among the three most disabling symptoms. SF-36 health status scores in the patients group were significantly lower than the obtained values of the controls and partially related to the FSS scores. Good internal consistency, significant test-retest reliability, and validity were obtained for the FSS.
Conclusions: Fatigue is a major symptom in patients with immune-mediated polyneuropathies and may persist for years after apparent recovery. The Fatigue Severity Scale seems appropriate for assessing fatigue in these patients, because good internal consistency, test-retest reliability, and validity were demonstrated.

Introduction

Guillain-Barré syndrome (GBS) and chronic inflammatory demyelinating polyneuropathy (CIDP) are immune-mediated diseases of the peripheral nervous system primarily characterised by symmetrical weakness, sensory disturbances, and reduction or loss of myotatic reflexes (1,2). Although the reported outcome percentages differ, approximately three-quarters of the patients with GBS or CIDP experience good physical recovery after adequate therapy (3). However, despite a good clinical improvement, we noticed that many of these patients still are restricted in their daily and social activities, sometimes consistently for many years. Moreover, these patients often address fatigue as the most important cause of their dysfunction. A decrement in their quality of life resulting from fatigue has also been claimed repeatedly.

Fatigue has been briefly addressed as a complaint in a limited number of GBS cases only (4,5). Because attention is primarily directed towards weakness and sensory disturbances in immune-mediated polyneuropathies, it is conceivable that fatigue may have been under-recognised by neurologists and rehabilitation physicians. To date, no cross-sectional study has compared the presence and severity of fatigue in these patients with healthy controls. Prompted by these considerations we investigated the presence and severity of fatigue in a cross-sectional group of patients with GBS, CIDP or a polyneuropathy associated with a monoclonal gammopathy of undetermined significance (MGUSP). It was argued that on the one hand these illnesses represent parts of a spectrum (3). On the other hand, this choice provided the opportunity to compare fatigue in patients with a more ongoing disease (CIDP/MGUSP) to those with a residual clinical condition due to GBS in the past. In addition, the internal consistency, test-retest reliability, and validity of the selected fatigue scale were examined.

Patients and methods

Participants

A group of 113 patients with a stable clinical condition (83 with GBS, 22 with CIDP, 8 with MGUSP) were recruited from the Rotterdam immune-mediated polyneuropathy databank and the Dutch GBS study group (table 1). A stable clinical condition was required to obtain the highest reliability for the selected fatigue scale. The selected patients still had residual symptoms or signs resulting from their illness, representing a broad range of disability (functional grading score [f-score] at least 1; see also assessment scales; 6). None of the patients have been treated with steroids or another immune-suppressive agent in the 6 months before the start of the study. Nine CIDP patients required interval treatment ranging from weeks to months, with intravenous immunoglobulin. With this therapy their clinical condition has been stable for more than 6 months. The patients with GBS and CIDP met the international criteria for their illness (1,2). The diagnosis MGUSP was established after excluding all possible underlying causes for the gammopathy and polyneuropathy (7). Six patients with MGUSP had an associated demyelinating polyneuropathy with minor concurrent axonal damage in three. An axonal polyneuropathy was diagnosed in the remaining two patients with MGUSP. Patients were excluded from participation if there was

any concomitant disease or use of medication that might cause chronic fatigue. All selected patients declared only to have experienced mild and transient fatigue prior to their illness. Controls, stratified for gender and age, were recruited from hospital personnel, companions (relatives, friends) of patients visiting our outpatient clinic, and volunteers unfamiliar with the study (see table 1). The controls declared to be healthy, free from any chronic medical condition, and were not taking medication that could contribute to fatigue.

Assessment scales

The *Fatigue severity scale* (FSS), a brief and simple self-assessed questionnaire, was selected and translated into Dutch before its use (8). The FSS is a nine-item questionnaire with answers ranging from 1 ("strongly disagree") to 7 ("strongly agree") for each inquiry. The mean score of the 9 inquiries ranges from 1 ("no signs of fatigue") to 7 ("most disabling fatigue"). Internal consistency, reliability, validity, and sensitivity of the FSS have been established in patients with systemic lupus erythematosus, multiple sclerosis, and Lyme disease, but never in patients with immune-mediated polyneuropathies (see appendix I, page 209; reference 8).

The *short-form fatigue scale (SFFS)*, with acceptable validity and reliability, was administered in order to evaluate the construct convergent validity of the FSS (9). The score of this scale ranges from 4 ("no signs of fatigue") to a maximum of 28 points ("most severe fatigue"). In contrast with the FSS, this scale does not relate fatigue to daily activities.

The *MRC sumscore* and *Functional grading scale (f-score)* were applied in the patient group to measure the physical condition of the patients at the impairment and disability levels (6). The MRC sumscore ranges from 0 ("paralysis") to 60 ("normal strength"). For each muscle group a standardised joint/limb position as well as the point at which counter-force is administered was pre-defined and taken when assessing muscle strength. The f-score of the patients included in this study ranges from 1 ("minor neurological symptoms or signs and able to run") to 4 ("bed or chair bound"). Good reliability and validity have been demonstrated for these two scales (6).

The *INCAT Sensory sumscore* ranges from 0 ("normal sensation") to 20 ("most severe sensory deficit") and is composed by the summation of the following sensation qualities: pinprick arm grade [range: 0 to 4] + vibration arm grade [range: 0 to 4] + pinprick leg grade [range: 0 to 4] + vibration leg grade [range: 0 to 4] + 2-point discrimination grade [range: 0 to 4]) (10). The vibration examination was performed using the recently validated graduated Rydel-Seiffer tuning fork and the obtained measures were compared with the reported vibration threshold normal values (11). Pinprick and vibration sense examination took place from distal to proximal, and only the highest extension of dysfunction of the most affected arm and leg was recorded for both qualities. The sites of examination with corresponding grades were defined as followed: normal (grade 0) or disturbed (grade 1) pinprick or vibration sense at the dorsum distal interphalangeal joint of the index finger or hallux; an abnormal sense at the ulnar styloid process or medial malleolus (grade 2), at the medial humerus epicondyle or patella (grade 3), and at acromioclavicular joint or anterior superior iliac spine (grade 4). For the 2-point discrimination quality, a sliding esthesiometer was used where the exact measurable distance in millimetres could be read on the instrument. This instrument was assessed in a 'static' manner at the ventral side, distal phalanx of the index

finger, and the corresponding grades were arbitrarily chosen (grade 0: ≤ 4mm; grade 1: 5 to 9mm; grade 2: 10 to 14mm; grade 3: 15 to 19mm; grade 4: ≥ 20mm) (10).
The *medical outcomes study 36-item short form health survey (SF-36)*, a generic health status questionnaire, consists of 36 items, assigned to the domains of physical functioning [10 items], role functioning-physical [4], role functioning-emotional [3], social functioning [2], body pain [2], mental health [5], vitality [4], general health perception [5], and change in health which is scored separately (12). The numbers of response categories per item range from two to six. Each domain has a scoring range from 0 to 100. A high score indicates better health or less body pain. The Dutch version we used was developed as a part of the IQOLA project, which aims to translate, validate, and norm the SF-36 in a range of languages and cultural settings (13).

Test procedure
All participants gave informed consent for the study. Participants were lucid and competent to answer the questionnaires to the best of their ability. All individuals received brief instructions on how to fill in the fatigue and health status forms. These questionnaires were answered in random order. Patients were examined at our outpatient clinic, and all scales were completed at study enrolment. The FSS was mailed 4 to 6 months later to the patients for a second assessment. The patients were also asked to report whether their clinical condition had improved, got worse, or remained the same since completing the first questionnaire. Patients whose health had remained the same were included in the FSS test-retest reliability analyses. The FSS was completed once by all healthy controls. SF-36 was self-administered by 59 controls, taking into account the age and sex distribution of the patients. The study took place between March 1997 and August 1998 and was part of a more comprehensive outcome assessment research in patients with immune-mediated polyneuropathies on behalf of the *INCAT* group.

Statistics
"Severe" fatigue was arbitrarily defined as a score ≥ FSS 95th percentile in healthy individuals. The obtained FSS scores in the patients and healthy controls were compared using the Mann-Whitney U-test (MWU-test) and univariate and multivariate quantile regression analyses (14). The median FSS values were compared within the categories (healthy controls, GBS, CIDP, and MGUSP patients) using design variables ("0 – 1") representing each category. The design variables were defined in an appropriate manner, making it possible to test (by the likelihood ratio test) the hypothesis that the median fatigue values were equal for the diagnosis CIDP and MGUSP, but unequal to the median fatigue values in GBS patients and healthy controls, separately. Box plots, with the upper and lower adjacent values defined according to Tukey, were applied to visualise some of these observations (also see legend to figure 1 for explanation of the various components of a box plot) (15).
The mean SF-36 subscale values were compared within the three diagnostic categories and between the whole patients group and healthy controls for each domain. One-way analysis-of-variance (Anova) with corrections according to Bonferroni multiple-comparison tests were applied.
The Cronbach's alpha coefficient was estimated for the FSS in the patients and healthy participants (16). The test-retest reliability for the recruited FSS values in the

polyneuropathy patients was quantified by estimation of the intraclass correlation coefficient using a one-way random effects analysis-of-variance model. The correlation between the FSS and other scales was analysed using Spearman's rank correlation test. All analyses were performed using Stata 5.0 for Windows 95 (Stata Statistical Software: Release 5.0. 702 University Drive East, College station, TX: Stata Corporation 1997). A value of $p \leq 0.05$ was considered statistically significant.

Table 1

Basic characteristics of the participants

Healthy controls (n=113)	
Sex, No (%)	
Females	54 (47.8%)
Males	59 (52.2%)
Mean age at start of the study, (SD) range [years]	54.2 (14.8) 18-83
Patients (n=113)	
Sex, No (%)	
Females	54 (47.8%)
Males	59 (52.2%)
Mean age at start of the study, (SD) range [years]	54.3 (15.1) 14-84
Number of patients per diagnosis	
GBS	83
CIDP	22
MGUSP	8
Median duration of symptoms till onset of study [years]	
Overall	5.1
GBS	5.2
CIDP	3.9
MGUSP	3.6
MRC sumscore distribution (score range: 0-60)	
< 49	25
50-59	66
60	22
f-score distribution (score range: 0-5; in current study: 1-4)	
1	51
2	41
3	14
4	7
Sensory sumscore distribution (score range: 0-20)	
0	21
1-9	73
10-19	19
20	0

Legend to table 1
The MRC sumscore ranges from 0 ("paralysis") to 60 ("normal strength"). The functional grading score (f-score) of the patients in the current study ranges from 1 ("minor neurological symptoms or signs and able to run") to 4 ("bed or chair bound"). The sensory sumscore ranges from 0 ("normal sensation") to 20 ("most severe sensory abnormalities").

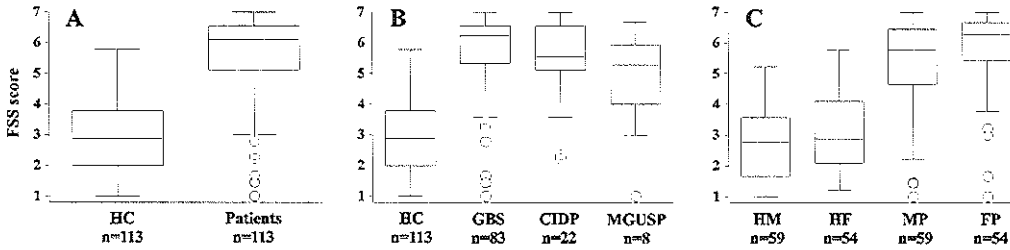
Results

Characteristics of participants. The characteristics of all participants are presented in table 1. One hundred and thirteen patients and 113 age and sex matched healthy individuals were enrolled. The CIDP patients were relatively younger (mean age 45.0 [SD 18.0] years)

compared with the GBS (mean age 56.1 [SD 13.5] years) and the MGUSP patients (mean age 66.1 [SD 8.1] years). One hundred and seven of the 113 patients (95%) completed the second FSS assessment. One hundred and five patients reported a stable clinical condition, while two reported an improved clinical condition at the second FSS assessment. These two were excluded from reliability analyses. Two patients died during the interval between the two measures, and another was hospitalised due to myocardial infarction. The remaining three patients did not return the second FSS assessment for various reasons. Fifty-nine healthy volunteers (29 females; 30 males; mean age 53.5 [SD 13.6] years) completed the SF-36 questionnaire.

Figure 1

Fatigue in patients with immune-mediated polyneuropathies and healthy controls (A), related to diagnosis (B) and sex (C)



Legend to figure 1

Box plots demonstrating the obtained 'Fatigue severity scale' (FSS) scores in the healthy controls versus patients group (A) related to diagnosis (B) and sex (C). HC = healthy controls, HM = healthy males, HF = healthy females, MP = male patients, FP = female patients. The patients were more fatigued than the healthy controls (Mann-Whitney U-test, whole patients group versus healthy controls: $p < 0.0001$). All diagnostic categories were more fatigued than the healthy controls (Likelihood Ratio-test combined with quantile regression analysis: $p < 0.0001$). The GBS patients had a higher median fatigue value compared with the other subgroups and healthy controls (Likelihood Ratio-test combined with quantile regression analysis: $p < 0.0001$ for the comparisons GBS versus CIDP + MGUSP, and GBS versus healthy controls). Females were more tired in both groups, but this was only significant in the patients group (Mann-Whitney U-test; in healthy controls: $p = 0.28$; in patients group: $p = 0.02$).

Explanation of the various components of the box plots: The median (50th percentile) FSS value corresponds with the horizontal line within the box. The box extends from the first quartile (25th percentile = x[25]) to the third quartile (75th percentile = x[75]), the so-called inter-quartile range (IQR). The vertical lines emerging from the box extend to the upper and lower adjacent values. The upper adjacent value is defined as the largest FSS value less than or equal to x[75] + 1.5 x IQR. The lower adjacent value is defined as the smallest FSS value greater than or equal to x[25] - 1.5 x IQR. Measured FSS values more extreme than the adjacent values are referred to as 'outside values' and are individually plotted (15).

FSS comparison between patients and controls. At entry, the median FSS value in the patients group was slightly higher (6.1) (corresponding mean value: 5.6 [SD 1.4]) compared with the median value at the second assessment (5.8) (corresponding mean value: 5.5 [SD 1.4]). Unless otherwise stated, further comparisons were made using only the FSS values

obtained at entry. The median FSS value in the patients group was higher compared with the result in healthy individuals (median fatigue value 2.9; corresponding mean value: 2.9 [SD 1.1]) (figure1A; MWU-test: $p < 0.0001$). By the Likelihood ratio test it was demonstrated that the median FSS values were equal for the CIDP and MGUSP patients but unequal to the median fatigue values in healthy controls and GBS patients. This test also revealed that the GBS patients had a higher median fatigue value (6.2) compared with the other subgroups and healthy controls (CIDP: 5.6; MGUSP: 5.3; healthy controls: 2.9; figure 1B; quantile regression analyses combined with Likelihood ratio test; $p < 0.0001$ for the comparisons GBS versus CIDP + MGUSP, and GBS versus healthy controls).
The FSS 95th percentile in the healthy controls was 5.0. "Severe" fatigue, arbitrarily defined as a FSS score ≥ than the FSS 95th percentile in the healthy controls, was demonstrated in 90 of 113 patients (80%). Remarkably, of the patients with a completely recovered general strength (MRC sumscore = 60) relatively more patients reported a fatigue score that fell within the severe fatigue range (19 of 22: 86%). Severe fatigue was also reported in 17 of the 21 (81%) patients with normal sensation (sensory sumscore = 0). Six of the 8 patients (75%) with both normal general strength and sensation experienced severe fatigue. Eighty percent of the patients scored 5 or more on the 7-point scale for question 8 ("Fatigue is among my three most disabling symptoms") compared with 12% (13/113) of the healthy controls.

Table 2

Correlations between the Fatigue Severity Scale (FSS) and SF-36 health status in patients with immune-mediated polyneuropathies

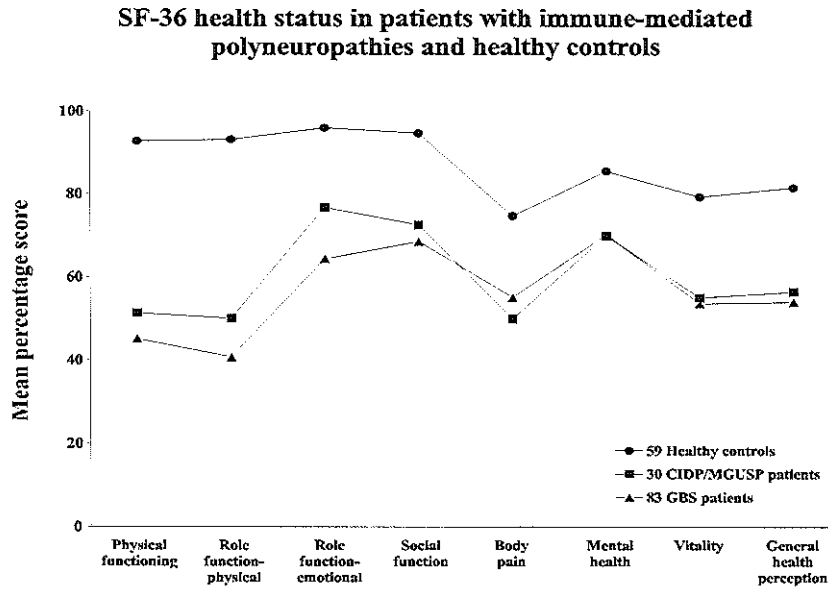
	GBS patients, n=83		CIDP/MGUSP patients, n=30	
	FSS	p-value	FSS	p-value
SF-36 domains	Spearman's rank correlation coefficient		Spearman's rank correlation coefficient	
Physical functioning	-0.21	NS	**-0.53**	**0.002**
Role functioning-physical	-0.17	NS	**-0.45**	**0.01**
Role functioning-emotional	**-0.22**	**0.04**	-0.33	NS
Social functioning	**-0.42**	**0.0001**	-0.24	NS
Body pain	**-0.23**	**0.03**	0.009	NS
Mental health	**-0.30**	**0.006**	-0.35	NS
Vitality	-0.57	< 0.0001	-0.50	0.005
General health perceptions	-0.40	0.0002	-0.65	0.0001

Legend to table 2
Fatigue in GBS patients was significantly associated with social and emotional domains, but not with the physical subscales, where as the CIDP/MGUSP patients demonstrated the contrary. NS = not significant.

FSS correlation with clinical parameters. General strength, as measured by the MRC sumscore, the f-score values, and sensory sumscore showed no significant association with the FSS values (Spearman's rank correlation test: $r = 0.07$ to 0.13; $p \geq 0.18$). Fatigue score was not significantly related to age and duration of symptoms in the patients group. Fatigue in females was consistently higher than in males, but this was only significant in the patients group (figure 1C; MWU-test in healthy controls: $p = 0.28$; in patients group: $p = 0.02$).

SF-36 evaluation and its association with FSS in patients. The SF-36 values showed no substantial difference between the three diagnostic categories (GBS, CIDP and MGUSP) within each domain (Anova with Bonferroni tests: all p-values ≥ 0.46). A separation was made between the GBS ("acute" subgroup; n = 83) and the CIDP/MGUSP patients ("chronic" subgroup; n = 30) based on these results combined with the FSS Likelihood ratio-test findings (see also results *FSS comparison between patients and controls*). For all

Figure 2



Legend to figure 2
Both patients' subgroups demonstrated a lower score for all SF-36 domains compared with the scores of the healthy controls (p < 0.0001 for all domains, except 'role functioning-emotional' in the CIDP/MGUSP patients versus controls: p = 0.03). No statistically significant difference was obtained at all domain levels between the GBS and the CIDP/MGUSP patients (all p-values ≥ 0.27). Analyses were performed using ANOVA with corrections according to the Bonferroni multiple-comparison tests.

SF-36 domains, both subgroups scored notable lower values compared with the scores obtained in healthy controls (figure 2; Anova with Bonferroni tests: p < 0.0001 for all domains, except for the dimension "role functioning-emotional" in the "chronic" subgroup versus controls: p = 0.03).
The correlations between the FSS and SF-36 domains for the two subgroups are presented in table 2. In both subgroups, FSS demonstrated the highest significant inverted association with the domains "general health perception" and "vitality". Fatigue scores in GBS patients were significantly associated with the socio-emotional domains, but not with the physical subscales. Conversely, fatigue values in the "chronic" subgroup were significantly related to the physical (but not the socio-emotional) dimensions (table 2, boxes).

Internal consistency, test-retest reliability, and validity evaluation of the FSS. A good internal consistency for the FSS was found at both assessments for the whole patient group (α = 0.93 and 0.95) and in healthy controls (α = 0.84). The FSS demonstrated good test-retest reliability and acceptable validity for the whole patient group when correlated with the SFFS and SF-36 "vitality" domain (r = 0.68 and r = -0.56, respectively; $p < 0.0001$; see also tables 2 and 3; mean SFFS was 17.6 (SD 7.2), range 4 to 28).

Table 3

Clinimetric evaluation of the Fatigue Severity Scale (FSS) in patients with immune-mediated polyneuropathies

Patients group (n=113)	Tests and results	p-value
Internal consistency	Cronbach's alpha	
FSS1	0.95	-
FSS2*	0.93	-
Reliability – test/retest	Reliability coefficient (Anova)	
FSS1 – FSS2	0.86	< 0.0001
Validity – construct convergent	Spearman's rank correlation coefficient	
FSS correlation with SFFS	0.68	< 0.0001
Validity – construct divergent	Spearman's rank correlation coefficient	
FSS correlation with SF-36 'vitality'	-0.56	< 0.0001

Legend to table 3
FSS1 = Fatigue Severity Scale measured at entry. FSS2 = Fatigue Severity Scale measured 4 to 6 months after FSS1. * = 105 of the 113 patients (93%) completed the second FSS assessment. SFFS = short-form fatigue scale (9).

Discussion

The findings in the current study demonstrate that fatigue is a prominent and highly disabling symptom in immune-mediated polyneuropathies. Four of five (80%) patients rated their fatigue as one of the three most disabling symptoms compared with only 12% in the healthy controls. Also, severe fatigue defined as a score equal to or higher than the FSS 95th percentile in the healthy controls, was present in 80% of the patients. These observations are consistent with reports on fatigue in other chronic conditions such as multiple sclerosis and systemic lupus erythematosus (17-20). The current study also demonstrated a difference with literature findings. The mean FSS score in our patients group was higher than the reported FSS values in various chronic medical conditions, except in patients with chronic fatigue syndrome (8,20-29; table 4). This was remarkable, since approximately two-thirds of the population at study consisted of patients with residual clinical conditions due to GBS many years (median: 5.2) before the start of this study. Surprisingly, the prevalence of fatigue was the highest among the GBS patients (figure 1B). Fatigue in these patients was independent of the time that elapsed since the acute phase of the GBS was experienced. Despite a good physical recovery, GBS can still be considered as a long-term event causing a "post-GBS fatigue syndrome" in most patients that may interfere with functionality. Such an "illness related fatigue syndrome" was also noted in the CIDP/MGUSP patients we investigated. Variables such as age, duration of the disease, MRC sumscore, f-score, and

sensory sumscore were not significantly associated with fatigue. Fatigue was only significantly related to sex in the patients group. Severe fatigue was also highly prevalent in patients with normal general strength or normal sensory modalities. These results clearly emphasise that fatigue should be considered as a serious, relatively independent and highly disabling entity in patients with immune-mediated polyneuropathies, irrespective of clinical variables or course of time.

The SF-36, a generic health status questionnaire, also clearly distinguished the patient group from healthy controls, but could not differentiate between the three diagnostic categories (GBS, CIDP and MGUSP) at all domain levels. The GBS ("acute") group and the "chronic" (CIDP/MGUSP) patients had significantly lower scores compared with the controls at all SF-36 levels (see figure 2). A significantly worse functional status has also been recently reported by Bernsen *et al.* using the sickness impact profile (SIP) health status in 123 patients who had GBS three to six years previously (30). These authors have also concluded that the psychosocial functioning of the patients was seriously affected, even when the patients have reached a complete physical recovery or showed only mild residual signs (30). The socio-emotional dimensions of the SF-36 applied in the current study are comparable with the psychosocial categories of the SIP and were fairly associated with fatigue in the GBS patients (31, table 2). Fatigue presumably contributes to a reduced psychosocial functioning in GBS patients. Contrary to these findings, fatigue was significantly associated with the SF-36 physical dimensions in patients with a "chronic" polyneuropathy. Perhaps these patients are more preoccupied with the potential threat of changes in their physical status making them more prone to relate disabling symptoms such as fatigue to their physical condition.

As a basis for their study, several papers addressing the influence of fatigue on health status in chronic disorders have stated that engagement in less physical activities and exercises may induce de-conditioning, which may explain fatigue leading to decrements in health status (19,32). Conversely, Pitetti *et al.* reported reduction in daily fatigue with improvements in activities of daily living following a supervised training program in a 54-year old man who had residual deficits for many years after having experienced GBS (33). An enhancement in functional capacity has also been reported in chronic relapsing GBS after low intensity aerobic exercise (34). Less engagement in physical activities was also noted in all patients in the current study compared with healthy controls (see dimensions "physical functioning" and "role-functioning physical" in figure 2). However, since fatigue was primarily related to the SF-36 physical domains only in the "chronic" group and not in the GBS patients, other explanatory factors leading to a reduced health status should also be considered.

Instruments for measuring outcome must be appropriate to the patient group being studied, efficient, easy to administer, and user friendly (35). Good reliability and validity are also necessary (35). In the current study, the FSS demonstrated to be brief and was easily self-administered with good test-retest reliability, and significant validity. The internal consistency of the FSS in the patients and healthy controls was higher than the recommended 0.7 score by Nunnally (16). Hence, the applicability of the FSS as an instrument for measuring fatigue is shown in patients with immune-mediated polyneuropathies. The responsiveness to changes in time of the FSS is currently being

Table 4

Fatigue severity scale (FSS) scores in current study compared with previously published FSS scores for other diagnostic groups and healthy controls

Groups studied: (reference)	Number of individuals studied	FSS Mean	FSS SD
Immune-mediated polyneuropathies			
Current study	113	5.6 (median: 6.1)	1.4
Multiple sclerosis			
Krupp *et al.*, 1989; (8)	25	4.8	1.3
Krupp *et al.*, 1993; (23)	57	5.1	1.4
Packer *et al.*, 1994; (24)	9	5.2	1.5
Bergamaschi *et al.*, 1997; (27)	100	4.1	?
Systemic lupus erythematosus			
Krupp *et al.*, 1989; (8)	29	4.7	1.5
Krupp *et al.*, 1990; (21)	59	4.6	1.5
Austin *et al.*, 1996; (25)	58	4.6	0.4
Wang *et al.*, 1998; (20)	100	5.1	3.1
Chronic fatigue syndrome			
Krupp *et al.*, 1993; (23)	72	6.1	0.8
Packer *et al.*, 1994; (24)	13	6.1	0.5
Postpolio syndrome			
Packer *et al.*, 1991; (22)	12	4.8	1.6
Packer *et al.*, 1994; (24)	28	5.1	1.7
Schanke, 1997; (28)	63	5.6	1.2
Lyme Borreliosis			
Ravdin *et al.*, 1996; (26)	21	4.7	1.7
Sleep disorders			
Lichstein *et al.*, 1997; (29)	206	4.8	1.4
Osteomyelitis			
Ravdin *et al.*, 1996; (26)	21	3.9	1.3
Healthy controls			
Current study	113	2.9	1.1
Krupp *et al.*, 1989; (8)	20	2.3	0.7
Krupp *et al.*, 1993; (23)	40	2.8	1.2
Packer *et al.*, 1994; (24)	11	2.2	1.1
Ravdin *et al.*, 1996; (26)	21	2.9	0.9

Legend to table 4
The mean FSS score of the patients at entry in the current study was higher than the literature values, except for the reported mean FSS scores in patients with chronic fatigue syndrome. The mean FSS score of the patients at the second assessment (4 to 6 months later) in the current study was 5.5 (SD 1.4).

evaluated in a longitudinal study including patients with recently diagnosed GBS and CIDP with changing clinical course, but has already been established in patients with MS and Lyme disease (8).

With respect to the aims of the current study, some methodological issues should be addressed. First, the selected patients had a stable clinical condition that was necessary for an optimal FSS reliability evaluation. However, a stable clinical condition was based on the subjective report by the patient rather than on the patient's actual clinical evaluation by an investigator. The latter would probably have resulted in a more accurate judgement of the clinical condition of the patient. Despite this, an acceptable FSS reliability value still was obtained. Second, it is unclear whether patients with changing clinical conditions would be more or less fatigued or whether the associations between FSS and SF-36 subscales would give different outcomes. Third, it is also unsettled whether fatigue would still not be related to the various clinical variables when assessed over time. Fourth, interpretation of the results

should be with some caution, since we did not systematically investigate or control for possible pathophysiological associated factors with fatigue, such as sleep disturbances, depression and inactivity leading to deconditioning (19-21,23,29,32-34). Despite these issues, the current observations still buttress our clinical experience that fatigue is a prominent complaint among patients with immune-mediated polyneuropathies.

In conclusion, fatigue is a highly prevalent and disabling symptom significantly associated with a reduced health status in patients with immune-mediated polyneuropathies. Further attention should be directed towards understanding and unravelling the pathophysiological mechanism of fatigue and introducing adequate therapeutic options for these patients.

References

1. Asbury AK, Cornblath DR. Assessment of current diagnostic criteria for Guillain-Barré Syndrome. Ann Neurol 1990;27(suppl):S21-S24.

2. Ad Hoc subcommittee of the American Academy of Neurology AIDS task force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

3. Hartung H-P, van der Meché FGA, Pollard JD. Guillain-Barré syndrome, CIDP and other chronic immune-mediated neuropathies. Curr Opin Neurol 1998;11:497-513.

4. Wessely S, Powell R. Fatigue syndromes: a comparison of chronic "postviral" fatigue with neuromuscular and affective disorders. J Neurol Neurosurg Psychiatry 1989;52:940-948.

5. Gregory RJ. Understanding and coping with neurological impairment. Rehabil Nurs 1995;20:74-83.

6. Kleyweg RP, van der Meché FGA, Schmitz PIM. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré syndrome. Muscle and Nerve 1991;14:1103-1109.

7. Miescher GC, Steck AJ. Paraproteinaemic neuropathies. Baillieres Clin Neurol 1996;5:219-232.

8. Krupp LB, LaRocca NG, Muir-Nash J, *et al.* The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 1989;46:1121-1123.

9. Alberts M, Smets EMA, Vercoulen JHMM, *et al.* 'Verkorte vermoeidheidsvragenlijst': een praktisch hulpmiddel bij het scoren van vermoeidheid. Ned Tijdschr Geneeskd 1997;141:1526-1530.

10. Merkies ISJ, Schmitz PIM, van der Meché FGA, *et al.* Psychometric evaluation of a new sensory scale in immune-mediated polyneuropathies. Neurology 2000;54:943-949.

11. Martina ISJ, van Koningsveld R, Schmitz PIM, *et al.* Measuring vibration threshold with a graduated tuning fork in normal ageing and in patients with polyneuropathy. J Neurol Neurosurg Psychiatry 1998;65:743-747.

12. McHorney CA, Ware JE jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993;31:247-263.

13. Aaronson NK, Acquadro C, Alonso J, *et al.* International quality of life assessment (IQOLA) project. Quality Life Res 1992;1:349-351.

14. Gould WW, Rogers WH. Quantile regression as an alternative to robust regression. 1994 proceedings of the statistical computing section. Alexandria, VA: American Statistical Association, 1994.

15. Tukey JW. Exploratory Data Analysis. Reading, MA: Addison-Wesley Publishing Company; 1977.

16. Nunnally JC. Psychometric theory. New York: McGraw Hill; 1978.

17. Murray TJ. Amantadine therapy for fatigue in multiple sclerosis. Can J Neurol Sci 1985;12:251-254.

18. Krupp LB, Alvarez LA, LaRocca NG, *et al.* Fatigue in multiple sclerosis. Arch Neurol 1988;45:435-437.

19. Robb-Nicholson LC, Daltroy L, Eaton H, *et al.* Effects of aerobic conditioning in lupus fatigue: a pilot study. Br J Rheumatol 1989;28:500-505.

20. Wang B, Gladman DD, Urowitz MR. Fatigue in lupus is not correlated with disease activity. J Rheumatol 1998; 25: 892-895.

21. Krupp LB, LaRocca NG, Muir J, *et al.* A study of fatigue in systemic lupus erythematosus. J Rheumatol 1990;17:1450-1452.

22. Packer TL, Martins I, Krefting L, *et al.* Post-polio sequelae. Activity and Post-polio fatigue. Orthopedics 1991;14:1223-1226.

23. Krupp LB, Jandorf L, Coyle PK, *et al.* Sleep disturbance in chronic fatigue syndrome. J Psychosom Res 1993;37:325-331.

24. Packer TL, Sauriol A, Brouwer B. Fatigue secondary to chronic illness: Postpolio syndrome, chronic fatigue syndrome, and multiple sclerosis. Arch Phys Med Rehabil 1994;75:1122-1126.

25. Austin JS, Maisiak RS, Macrina DM, *et al.* Health outcome improvements in patients with systemic lupus erythematosus using two telephone counselling interventions. Arthritis Care Res 1996;9:391-399.

26. Ravdin LD, Hilton E, Primeau M, *et al.* Memory functioning in Lyme Borreliosis. J Clin Psychiatry 1996;57:282-286.

27. Bergamaschi R, Romani A, Versino M, *et al.* Clinical aspects of fatigue in multiple sclerosis. Funct Neurol 1997;12:247-251.

28. Schanke A-K. Psychological distress, social support and coping behaviour among polio survivors: a 5-year perspective on 63 polio patients. Disabil Rehabil 1997;19:108-116.

29. Lichstein KL, Means MK, Noe SL, *et al.* Fatigue and sleep disorders. Behav Res Ther 1997;35:733-740.

30. Bernsen RAJAM, Jacobs HM, de Jager AEJ, *et al.* Residual health status after Guillain-Barré syndrome. J Neurol Neurosurg Psychiatry 1997;62:637-640.

31. Katz JN, Larson MG, Phillips CB, *et al.* Comparative measurement sensitivity of short and longer health status instruments. Med Care 1992;30:917-925.

32. Petajan JH, Gappmaier E, White AT, *et al.* Impact of aerobic training on fitness and quality of life in multiple sclerosis. Ann Neurol 1996;39:432-441.

33. Pitetti KH, Barrett PJ, Abbas D. Endurance exercise training in Guillain-Barré syndrome. Arch Phys Med Rehabil 1993;74:761-765.

34. Karper WB. Effects of low-intensity aerobic exercise on one subject with chronic-relapsing Guillain-Barré syndrome. Rehabil Nurs 1991;16:96-98.

35. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. J Neurol Neurosurg Psychiatry 1996;60:127-130.

Chapter 10

Quality of life complements traditional outcome measures in immune-mediated polyneuropathies

Merkies ISJ[1], Schmitz PIM[2], van der Meché FGA[1], Samijn JPA[1], van Doorn PA[1], for the Inflammatory Neuropathy Cause and Treatment (INCAT) group[3]

From the Department of [1]Neurology, [2]Department of Statistics - Daniel den Hoed Cancer Centre, Erasmus University Medical Centre Rotterdam, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands. [3]See Chapter 2 for INCAT-members.

Submitted for publication

Abstract

Background: Quality of life scales assess outcome from patients' perspectives. It was argued that these scales complement traditional outcome measures, thereby providing a more comprehensive view of how illnesses affect patients' lives.
Objectives: Quality of life was, therefore, examined in immune-mediated polyneuropathies using the medical-outcome-study 36-items (SF-36). Its psychometric requirements were also analysed.
Methods: SF-36 and three other measures (MRC sumscore, Sensory sumscore, Hughes's functional scale) were assessed in 114 stable patients (83 with Guillain-Barré syndrome (GBS), 23 with chronic inflammatory demyelinating polyneuropathy (CIDP), 8 patients with a gammopathy related polyneuropathy) and serially in 20 patients with recently diagnosed GBS (n=7) or CIDP (n=13) with changing conditions. The SF-36 values were compared with reported healthy Dutch community scores (controls). The SF-36 validity and reliability were examined by correlation and regression studies with the other measures and by calculating its internal consistency. The standardised-response-mean technique was applied to determine its responsiveness.
Results: In the stable group, all SF-36 scores were substantially lower (indicating worse clinical condition) compared with controls' ($p<0.0001$). Improvement in the longitudinal group resulted in a gradual shift of all SF-36 scores towards normal values. Acceptable validity and internal consistency values and high standardised-response-mean scores were demonstrated for the SF-36. The MRC sumscore and sensory sumscore explained SF-36 values only partially.
Conclusion: The psychometric requirements of the SF-36 are demonstrated in immune-mediated polyneuropathies. This generic health status complemented traditional strength and sensory measures and appears to be a potentially valuable instrument for measuring quality of life in these conditions.

Introduction

In the last decade, an exponential increase has been witnessed regarding the use of health-related quality of life measures in clinical studies. This increase was driven by the concept that traditional clinical and laboratory measures needed to be complemented by measures that focus on the patients' concerns, thereby striving for a better understanding on how patients' lives are being affected by their disorders and to identify more appropriate forms of health care (1).
Particularly in diseases with a long-term effect on functionality such as immune-mediated polyneuropathies, more attention should be paid to the impact of illness and its treatment on functional, emotional, and social well-being of patients (1). However, despite this growth only a limited number of clinical reports included a health status measure in the evaluation of patients with Guillain-Barré syndrome (GBS), chronic inflammatory demyelinating polyneuropathy (CIDP), or a polyneuropathy associated with a monoclonal gammopathy of undetermined significance (MGUSP) (2-5). Moreover, none of the applied quality of life measures was thoroughly evaluated in terms of being valid, reliable, and responsive to changes over time before its general use in these illnesses (6).
Prompted by these findings, we investigated quality of life in patients with GBS, CIDP and MGUSP using the medical outcome study 36-items short-form health status (SF-36) and the obtained values were compared with the published normal data for the healthy Dutch community (7-10). This generic health status was chosen because of its brevity and extensive use in clinical studies. In addition, the association was examined between the SF-36 and variables like age, duration of symptoms, general strength, sensation, and functional ability. The discriminatory ability of the SF-36 between groups of patients with various degrees of functional ability (discriminatory ability) and its reliability were also analysed (6). Finally, we investigated whether immune-mediated polyneuropathies would lead not only to physical deficit, but also to mental disturbances and whether extensive medical guidance would provide a better quality of life over time.

Patients and methods

Participants
A cross-sectional group of 114 Dutch patients (83 with GBS, 23 with CIDP, 8 with MGUSP) with a stable clinical condition were recruited from the Rotterdam immune-mediated polyneuropathy databank and the Dutch GBS study group (stable group). These patients came from all Dutch districts and were treated in various hospitals through out The Netherlands. The patients did not have any known concomitant disease that might influence quality of life. These patients still had residual symptoms or signs due to their illness, representing a broad range of disability: functional grading scale (f-score) ranging from 1 to 4 (see also assessment scales and reference 11). Nine CIDP patients required interval treatment ranging from weeks to months, with intravenous immunoglobulin (IVIg). With this therapy their clinical condition has been stable for more than 6 months. Six patients with MGUSP had an associated demyelinating polyneuropathy with minor concurrent axonal damage in three. An axonal polyneuropathy was diagnosed in the remaining two patients with MGUSP.

Twenty consecutive patients at the university hospital Rotterdam with newly diagnosed GBS (n = 7) or CIDP (n = 13) and changing clinical conditions were enrolled to investigate whether the SF-36 captured longitudinal changes in these disorders (longitudinal group). These patients were also free from any concomitant disease. The baseline f-score values in these patients ranged from 1 to 5 (see also assessment scales and reference 11). During the recruitment period, no patients with newly diagnosed MGUSP were seen at our department. All GBS and CIDP patients met the international criteria for their illness (12,13). The diagnosis MGUSP was established after excluding all possible causes for the gammopathy and polyneuropathy (14).
The reported mean (standard deviation [SD]) SF-36 domain scores and summary values for a random, nationwide sample of 1742 healthy individuals in The Netherlands (976 men, 766 women; mean age 47.6 [18.0], range 16-94 years) was used for the comparison studies with the SF-36 values in the selected patients groups (9,10).

Assessment scales
The *medical outcome study 36-items short form health survey* (*SF-36*), a generic health status questionnaire, consists of 36 items, assigned to the domains of physical functioning [10 items], role functioning-physical [4], role functioning-emotional [3], social functioning [2], body pain [2], mental health [5], vitality [4], general health perception [5], and change in health which is scored separately (see appendix IV, page 217-221 and reference 7). The numbers of response categories per item range from 2 to 6. Each domain has a scoring range from 0 to 100. A high score indicates better health or less body pain. The Dutch version of the SF-36 was used (9). The corresponding physical component summary (PCS) and mental component summary (MCS) values were also calculated (8).
The *Medical Research Council (MRC) sumscore* is a summation of the MRC grades given in full numbers of the following muscle pairs: upper arm abductors, elbow flexors, wrist extensors, hip flexors, knee extensors, and foot dorsal flexors (11). The MRC sumscore ranges from 0 ("total paralysis") to 60 ("normal strength") (11).
The "*INCAT*" *sensory sumscore* (ISS) was applied to measure general sensory deficit (15). The ISS comprises vibration and pinprick sense plus a 2-point discrimination value and ranges from 0 ("normal sensation") to 20 ("most severe sensory deficit") (15).
The *Hughes' functional grading scale* (*f-score*) assesses the functional ability of the patients (11). The f-score of the patients included in this study ranged from 0 to 5. An f-score of 0 indicates normal health (no symptoms or signs); 1, denotes having minor neurological symptoms or signs and able to run; 2, able to walk at least 10 meters, but unable to run; 3, able to walk 10 meters with a walker or support; 4, bedridden or chair bound (unable to walk 10 meters with a walker or support); 5, requiring artificial ventilation for at least part of the day (11). Good psychometric properties were demonstrated for all selected scales (8-11,15). However, SF-36 has not been evaluated in terms of its validity, reliability, and responsiveness in immune-mediated polyneuropathies.

Test procedures and treatment
All participants gave informed consent before the study. All measures were obtained at our outpatient clinic. The patients received instructions on how to fill in the SF-36. The time to complete the SF-36 was recorded in fifty participants. The assessments were performed in a random order. For the assessment of strength, a pre-defined standardised joint and limb

position as well as the point at which counter-force was administered was taken for each muscle group. Sensory modalities were examined in triplicate according to the described standard procedures with the patient lying in supine position (15). In the stable group of patients, all measures were examined once. Before leaving, all SF-36 items were checked on possible missing values and if necessary patients were asked to complete the missing questions.

The longitudinal group of patients was examined by the same clinician (IM) and the MRC sumscore, ISS, and f-score were assessed randomly at the weeks 0, 2, 4, 8, 12, 16, 21, 26, 32, 40, and 52 of follow-up in each patient. If necessary, additional clinical investigations were performed. The SF-36 was concurrently assessed at the weeks 0, 4, 12, 26, 32, 40, and 52. All SF-36 domains and summary measures were scored as recommended by the SF-36 developers (7,8). Patients were also allowed to contact us between clinical visits for any question possible related to their illness. Family members of the patients were also stimulated to participate in the whole process of dealing with the consequences of the illness.

All but one mildly affected patient were initially treated with IVIg. IVIg demonstrated to be efficacious in GBS and CIDP (16-19). One IVIg non-responsive CIDP patient received additional treatment with prednisone (20). The study took place between March 1997 and September 1999 and was part of a more comprehensive outcome assessment research in patients with immune-mediated polyneuropathies on behalf of the *INCAT*-group.

Statistics

Cross-sectional study. In the stable group, mean SF-36 subscales' and summary (PCS and MCS) values were compared with the reported mean normal values for the Dutch community (Student's t-test for two independent groups) (9,10). Also, the discriminatory capacity of the SF-36 was examined in two groups of patients with different degrees of disability (subgroup 1: able to walk independently [f-scores 1 and 2] versus subgroup 2: needing support to walk or unable to walk [f-scores 3 and 4]). The possible correlation between MRC sumscore, ISS values, and f-score values with SF-36 domains and summary measures were calculated using Spearman's rank correlation coefficient. Reliability of the SF-36 subscales and summary measures was estimated by calculating the internal consistency (Cronbach's α). According to Nunnally, a $\alpha \geq 0.7$ was considered as good reliability (21). The Cronbach's α values for the PCS and MCS were calculated using a covariance matrix for the eight SF-36 dimensions as suggested by Ware and associates (8).

Univariate linear regression analyses were also performed to determine the impact of age, duration of symptoms till onset of the study, general strength, and sensation (explanatory variables) on the PCS and MCS values separately (dependent variables). The PCS values had a Gaussian pattern. A transformation (quadratic form) of the MCS scores was needed to obtain a normal distribution pattern prior to the regression studies. Through systematic construction and evaluation of graphs, we strived for the best fit between the explanatory and dependent variable using restricted cubic spline function on the independent variables in the regression studies (22). The strength of association between these variables was presented as R^2: the fraction of variance explained by the independent variable from the regression model.

Longitudinal study. In the longitudinally examined patients, mean SF-36 subscales' and summary scores were compared also at various arbitrarily chosen occasions of follow-up

(weeks 0, 12, 26, 52) with the reported SF-36 normal Dutch community values (Student's t-test) (9,10). Moreover, random effects univariate linear regression analyses were performed to investigate the impact of general strength, sensation, and functional ability on the SF-36 scores, taking into account the correlation of the data caused by the longitudinal structure. The latter was achieved using the program "xtreg" in STATA 6.0 for Windows 95 which is based upon a cross-sectional time-series regression model as described by Dwyer and Feinleib (StataCorp. 1997. Stata Statistical Software: Release 6.0. College Station, TX: Stata Corporation; see also reference 23). Before applying "xtreg", the distribution patterns of all variables were examined and where necessary and possible a transformation of the variable(s) took place to ascertain a normal distribution. Also, graphics of each clinical (explanatory) variable and a SF-36 domain or summary measure (dependent variables) were constructed and examined. If necessary, additional transformation of the variables was performed (e.g. logarithmic) to obtain the best fit for the regression function. The strength of association between these variables was presented as R^2: the fraction of variance explained by the independent variable + 'time'-factor from the regression model.
SF-36 responsiveness was estimated at 6 and 12 months of follow-up using the standardised response mean score (SRM; 24). SRM is equal to the mean change in scores divided by the standard deviation of change in scores ($SRM = (\mu i - \mu o)/SD(\mu i - \mu o)$; μi = mean scale value of the longitudinally examined group at week = i; μo = mean scale value at week = 0 [entry]) (24). A value between 0.5 and 0.8 is considered moderate, and 0.8 or greater as high responsiveness (24,25). All analyses were performed using Stata 6.0 for Windows 95. A value of $p < 0.05$ was considered statistically significant.

Results

Patients' characteristics and general aspects. The basic characteristics of all patients are presented in Table 1. In the stable group, fourteen patients required assistance or a device to walk short distances (f-score = 3) and eight patients were bed bound (f-score = 4) (22/114 = 19%; 17 GBS, 3 CIDP, 2 MGUSP). The remaining 92 patients (81%) could walk independently (f-score ≤ 2; 66 GBS, 20 CIDP, 6 MGUSP; Table 1). At study entry, thirteen longitudinally examined patients were unable to walk independently (4 GBS, 5 CIDP) or were bed bound (2 GBS, 2 CIDP), one requiring artificial ventilation. The remaining 7 patients (1 GBS, 6 CIDP) could walk independently.
According to all patients, the SF-36 was easily administered and took 9 (SD 2) minutes for completion. All SF-36 items were completed. The longitudinally examined patients contacted us with questions (usually once a week by telephone), especially within the first three months of their illness. All patients reported having the opportunity to contact us as a big support in managing the consequences of their illness.
SF-36 scores in stable patients. The mean SF-36 subscales' values with corresponding mean PCS and MCS scores for the stable group of 114 patients were notably lower compared with the reported mean normal values (Figure 1A and 1B; t-test: $p < 0.0001$). Differences between independent ambulatory patients (n = 92; f-scores ≤ 2) and those patients who needed support to walk or could not walk (n = 22; f-scores = 3 and 4) are presented in Figures 1C and 1D. These differences were primarily reflected in the more physically

Table 1

Basic characteristics of patients with immune-mediated polyneuropathies	
Stable group of patients (n=114; GBS 83, CIDP 23, MGUSP 8)	
Sex, No (%)	
Females	54 (474%)
Males	60 (53%)
Mean age at start of the study, (SD) range [years]	55.0 (15.2) 14 – 84
Mean duration of symptoms till onset of study [years]	6.8 (6.7) 0.7 – 28
MRC sumscore	53.2 (7.9) 16 – 60
Sensory sumscore	4.4 (4.1) 0 – 15
Functional grading scale (f-score) at entry	
f-score = 1	51 (45%)
f-score = 2	41 (36%)
f-score = 3	14 (12%)
f-score = 4	8 (7%)
Longitudinal group of patients (n=20; GBS 7, CIDP 13)	
Sex, No (%)	
Females	8 (40%)
Males	12 (60%)
Mean age at start of the study, (SD) range [years]	45.6 (18.5) 15 – 70

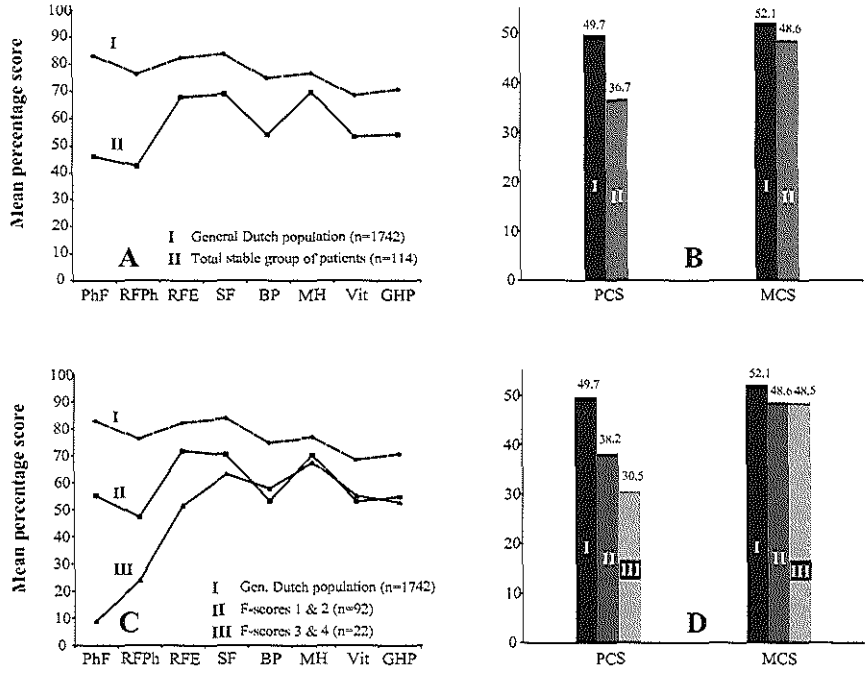
Legend to Table 1
GBS = Guillian-Barré Syndrome; CIDP = chronic inflammatory demyelinating polyneuropathy; MGUSP = monoclonal gammopathy of undetermined significance associated polyneuropathy.
MRC sumscore range: 0 ("total paralysis") to 60 ("normal strength"); "INCAT" sensory sumscore range: 0 ("normal sensation") to 20 ("most severe sensory deficit"); Hughes' functional grading scale (f-score) range: 0 ("no symptoms or signs") to 5 ("requiring artificial ventilation for at least part of the day") (11,15).

focused domains ("physical-functioning" and "role-functioning physical"; t-test: $p \leq 0.002$; for the dimension "role-functioning emotional": $p = 0.02$; for the remaining subscales: $p \geq 0.1$). The PCS scores clearly distinguished between the two disability subgroups, demonstrating a lower mean PCS value for the more disabled subgroup (Figure 1D; t-test: $p = 0.0007$). No differences between these subgroups were seen for the MCS values (Figure 1D; t-test: $p = 0.5$).

SF-36 values in longitudinally examined patients. Eight women and twelve men were examined longitudinally. Two hundred and one visits were completed in these patients. SF-36 was concurrently assessed in 147 occasions of these 201 visits (5 - 10 visits in each patient). The follow-up period ranged from 40 – 58 (mean 52) weeks. Nineteen patients completed a one-year follow-up. All patients experienced general loss of strength, sensory disturbances, and reported functional deficit. Good clinical improvement was noted during follow-up in all patients. The GBS patients (n = 7; 6 were treated with IVIg; 1 patient received no treatment) did not show any deterioration during follow-up. Twelve IVIg responsive CIDP patients needed interval therapy with IVIg (1-2 days 0.4 grams/kg/day; intervals: 3 - 21 weeks) to maintain the achieved improvement. One CIDP patient did not respond on IVIg therapy and was treated with prednisone (initial dose 100 milligrams/day for 4 weeks, tapered down over 5 months to 30 milligrams every other day). Eventually, all patients demonstrated during follow-up a general increase in quality of life as shown in Figure 2. To provide clarity, only the SF-36 values at the weeks 0, 12, 26, and 52 of follow-up were graphically presented. As can be seen, the SF-36 adequately captured improvement in these patients, demonstrating a gradual shift of all scores in the whole group towards

Figure 1

Comparison of mean SF-36 health status profile in a stable group of patients with immune-mediated polyneuropathies versus mean values for the healthy Dutch community



Legend to Figure 1

PhF = physical functioning; RFPh = role-functioning physical; RFE = role-functioning emotional; SF = social functioning; BP = body pain; MH = mental health; Vit = vitality; GHP = general health perception; PCS = physical component summary score; MCS = mental component summary score. Figure 1A: mean domains' values are presented for the whole stable group; Figure 1B: mean component summary scores are presented for the whole stable group. Mean dimensions and summary measures were notably lower compared with corresponding values in healthy controls (t-test: $p < 0.0001$ for all comparisons). Figure 1C: mean domains' values are presented for the two disability subgroups (: f-score ≤ 2: able to walk independent versus f-scores 3 + 4: able to walk with support or bed bound); Figure 1D: mean component summary scores are presented for the two disability subgroups. Differences between these two subgroups were especially reflected in the more physically focused domains ("physical-functioning" and "role-functioning physical"; t-test: $p \leq 0.002$; for the domain "role-functioning emotional": $p = 0.02$; for the remaining dimensions: $p \geq 0.1$). The reported mean (SD) SF-36 domain scores with corresponding summary values for a nationwide sample of 1742 healthy individuals in The Netherlands was used for the comparison studies (9,10).

normal values (Figure 2). At entry, all dimensions and summary measures were notably lower compared with the corresponding normal values (t-test: $p < 0.0001$ for all comparisons). Gradual but substantial improvements were seen during treatment and at one year of follow-up only the domain "physical-functioning" and PCS differed significantly from the corresponding normal values (Figure 2). The MRC sumscore and ISS also captured

Table 2

MRC sumscore, *INCAT* Sensory sumscore (ISS), and Hughes' functional grading (f-score) values in longitudinally examined patients with immune-mediated polyneuropathies

1	At entry	At week 12	At week 26	At week 52*
MRC-ss; mean (SD), range	47.6 (10.4), 13 – 58	53.9 (9.7), 28 – 60	56.4 (7.2), 30 – 60	57.4 (6), 37 – 60
ISS; mean (SD), range	8.8 (4.6), 1 – 18	5.4 (5.3), 0 – 14	3.7 (4.2), 0 – 12	4.0 (5.3), 0 – 16
f-score 0 – 2	7	16	18	17
3 – 5	13	4	2	2

Legend to Table 2
MRC-ss = MRC sumscore. * = one patient did not complete one year of follow-up. Note: a lower ISS or f-score connotes less impairment and disability, respectively. Disability grades: f-score ≤ 2: able to walk independent; f-scores 3 + 4: able to walk with support or bed bound. MRC sumscore range: 0 ("total paralysis") to 60 ("normal strength"); "INCAT" sensory sumscore range: 0 ("normal sensation") to 20 ("most severe sensory deficit"); Hughes' functional grading scale (f-score) range: 0 ("no symptoms or signs") to 5 ("requiring artificial ventilation for at least part of the day") (11,15).

improvement over time (Table 2). Moreover, lower f-score values were seen in most patients during follow-up, hence representing a better functional status (Table 2).

Clinical variables and their impact on SF-36 PCS and MCS scores. Age was not significantly related to the PCS and MCS values in both patients' groups. Also, duration of symptoms was not significantly correlated with the summary values in the stable group.

General strength and sensory disturbances explained partially the obtained PCS values (in the stable group: MRC sumscore: fraction explained variation: $R^2 = 0.12$; ISS: $R^2 = 0.16$; in longitudinal group: MRC sumscore: $R^2 = 0.17$; ISS: $R^2 = 0.36$; $p<0.0001$ for all regressions). MRC sumscore and ISS values were not significantly related to the MCS scores in the stable group. In the longitudinally examined patients, both variables had a significant impact on MCS (MRC sumscore explaining 11% and sensory sumscore 13% of MCS values; $p<0.0001$ for all regressions).

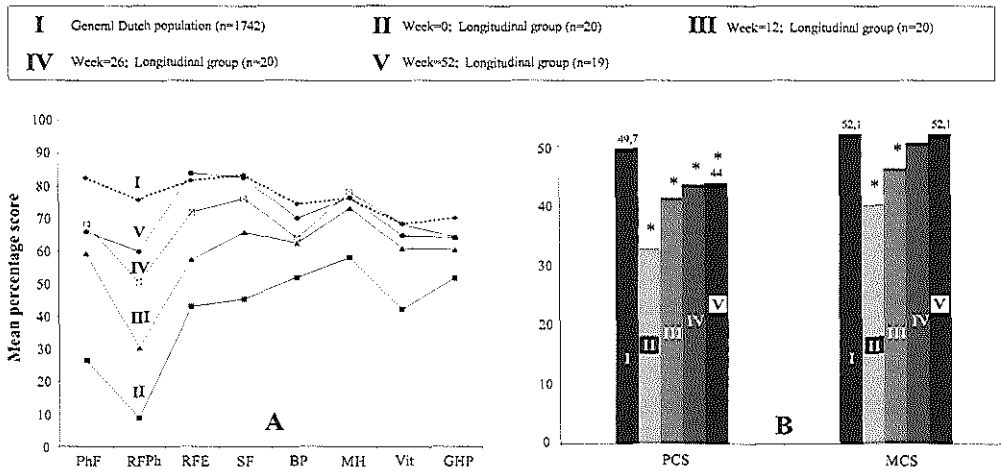
Validity, Reliability, and Responsiveness of the SF-36. The psychometric requirements for the SF-36 are presented in Table 3. Poor to good correlation values were obtained between the SF-36 and other clinical measures. The highest associations were seen between the physically oriented SF-36 entities, and the MRC sumscore and f-score values (Table 3). Good internal consistency values were obtained for all SF-36 dimensions and summary measures. All dimensions and summary measures, except "body pain" and "general health perception", had SRM values ≥ 0.8, indicating high responsiveness. Moderate SRM values (0.4 –0.6) were obtained for "body pain" and "general health perception" (Table 3).

Discussion

The current study demonstrates the clinical applicability of the SF-36 generic health status in patients with immune-mediated polyneuropathies. Patients with a stable GBS, CIDP, or MGUSP for many years had significantly lower SF-36 scores at all levels, indicating a worse clinical state. Particularly the more physically oriented entities demonstrated lower values

Figure 2

Mean SF-36 changes in a longitudinal group of patients with immune-mediated polyneuropathies receiving maintenance therapy versus mean values for the healthy Dutch community



Legend to Figure 2

Figure 2A: mean domains' values are presented for the longitudinal group; Figure 2B: mean component summary measures are presented for the longitudinal group. PhF = physical functioning; RFPh = role-functioning physical; RFE = role-functioning emotional; SF = social functioning; BP = body pain; MH = mental health; Vit = vitality; GHP = general health perception; PCS = physical component summary score; MCS = mental component summary score. Twenty patients were examined longitudinally. One hundred and forty seven SF-36 measures were completed. There was a gradual shift of all SF-36 domains and summary measures towards normal values (II $\rightarrow$ V). * $p \leq 0.01$ for comparison with healthy controls (t-test). The reported mean (SD) SF-36 domain scores and summary measures for a random sample of 1742 healthy individuals in The Netherlands was used for the comparison studies (9,10).

compared with the reported normal values for the Dutch community (Figure 1). Patients who were more disabled had lower scores on the physical measures compared with the less disabled ones. Similar SF-36 discriminatory findings were demonstrated in a cross-sectional group of patients with systemic lupus erythematosus (26). Surprisingly, in the current study, the mental state did not decrease with increasing disability and was almost equivalent to the controls. An unaltered mental state, scored by the SF-36, was also reported in patients with various forms of chronic peripheral neuropathies (5). Although statistically significant, the differences in the MCS values between the stable patients and controls were minor and most probably clinically not relevant. In the longitudinally examined patients, the more mentally oriented subscales (mental health, social function, body pain) and the MCS values reached earlier normal values compared with the more physical SF-36 measures (physical function, role-function physical, PCS) (Figure 2). Comparable results were reported in patients with rheumatoid arthritis who demonstrated an improvement in psychosocial functionality as the duration of the disease increased (27,28). It was suggested that in general patients alter their functional expectations over time and learn to cope with their limitations (28,29). Mental health and subjective well-being were also by far the least affected in patients with various

Table 3

Validity, Reliability, and Responsiveness of SF-36 health status in patients with immune-mediated polyneuropathies

	Validity			Reliability	Responsiveness	
	In stable group: Spearman's rank correlation coefficient (r) In longitudinal group: Random effects linear regressions (R)			Internal consistency (Cronbach's α)	Standardised Response mean statistic at:	
					6 months	12 months
Stable group of patients (n=114)						
SF-36 domains + summary scores	MRC Sumscore	INCAT Sensory sumscore	Functional grading scale (f-score)			
PhF	0.60**	-0.41**	-0.76**	0.95	-	-
RFPh	0.20	-0.13	-0.27*	0.85	-	-
RFE	0.01	0.09	0.01	0.83	-	-
SF	0.17	-0.09	-0.17	0.77	-	-
BP	-0.03	-0.23*	-0.08	0.72	-	-
MH	0.04	0.03	-0.04	0.87	-	-
Vit	-0.06	-0.07	-0.01	0.80	-	-
GHP	0.10	-0.08	-0.18	0.80	-	-
PCS	0.37**	-0.42**	-0.56**	0.84	-	-
MCS	-0.09	0.17	0.16	0.90	-	-
Longitudinal group of patients (n=20; 147 visits)						
PhF	0.81**	-0.45**	-0.85**	-	1.4	1.3
RFPh	0.32**	-0.53**	-0.52**	-	1.1	1.0
RFE	0.32**	-0.41**	-0.43**	-	0.8	1.1
SF	0.68**	-0.38**	-0.68**	-	0.8	0.9
BP	0.06	-0.40**	-0.24*	-	0.4	0.6
MH	0.37**	-0.48**	-0.43**	-	1.1	1.0
Vit	0.28**	-0.40**	-0.35**	-	1.1	0.9
GHP	0.28**	-0.51**	-0.44**	-	0.5	0.4
PCS	0.41**	-0.60**	-0.66**	-	1.0	0.9
MCS	0.29**	-0.36**	-0.43**	-	1.1	1.2

Legend to Table 3
PhF = physical functioning; RFPh = role-functioning physical; RFE = role-functioning emotional; SF = social functioning; BP = body pain; MH = mental health; Vit = vitality; GHP = general health perception; PCS = physical component summary score; MCS = mental component summary score. * $p \leq 0.01$; ** $p \leq 0.0001$.

chronic medical conditions (30). There are, however, exceptions to this statement, since patients with epilepsy or aids may have an altered emotional well-being and mental functioning over time (31,32).

In the longitudinally examined patients, clinical improvement over time was also demonstrated by a gradual shift of all SF-36 values towards the normal values (Figure 2). The pattern of changes of the SF-36 was more visible in some domains than in others and therefore it is unlikely that these changes were solely due to regression to the mean. Our results emphasize that GBS and CIDP have pervasive impacts on patients' function, both physical and psychosocial. The SF-36, therefore, complements the traditional viewing of symptoms, signs, and laboratory studies in these conditions and facilitates the evaluation of not only physical, but also mental functioning, hereby increasing the physician's awareness of the magnitude and extent of their patients' functional limitations. This knowledge may lead to early educational interventions of patients and their family about the general course and prognosis of GBS or CIDP, management of possible fear, anxiety, and other psychological discomfort (33). The longitudinally examined patients experienced an

extensive follow-up regime and guidance that most probably contributed to a better outcome. It was argued that a supportive group of individuals surrounding a patient plays an important role in the promotion of health and well-being and therefore ultimately in the evaluation of quality of life of the patient (34). It should be noted, however, that the quality of life improvement seen in the longitudinally examined patients was most probably related to a mixture of supportive care and pharmacological therapy with 'known' efficacy (16-20).

Quality of life was also examined using the SF-36 dimension and summary scores in a randomised controlled trial evaluating the effect of a home exercise program in a small sample of patients with chronic peripheral neuropathies (5). The findings in this paper were congruent with our results. The more physically oriented dimensions plus the PCS scores were more altered compared with the mentally oriented entities. Also, all SF-36 scores except for the MCS values were lower than scores previously described for the general population (5). Another paper investigated quality of life using the Sickness Impact Profile (SIP) generic health survey in a cross-sectional group of 123 patients who experienced GBS 3 to 6 years earlier (2). In accordance with our findings, the physical and psychosocial dimensions of the SIP were clearly lower compared with the healthy controls'. Lower physical functioning with increasing disability was also noted (2).

General strength explained 12-17% and sensory deficit 16-36% of the SF-36 physical summary scores, which is comparable with earlier findings in various forms of polyneuropathy (4). In the longitudinally examined patients, general strength and sensory deficit explained only a small portion of the mental summary scores. These findings implicate that the SF-36 could be considered as an adjunct outcome measure for future studies, complementing traditional strength and sensory scales. Moreover, in understanding what causes a decrement in quality of life in these conditions, other explanatory variables need to be considered. Variables like depression and fatigue have been advocated as potential contributors to a decrement in quality of life (33,35).

In the longitudinally examined patients, the SF-36 "mental health" subscale values, containing various aspects of depression, and the MCS scores during the first 6 months of follow-up were equivalent to the scores that were reported in mildly depressed patients (36). Since a significant correlation has been demonstrated between the SF-36 "mental health" domain, MCS, and a depression scale in patients suffering from depression, it is conceivable that depression may have contributed to a psychosocial dysfunction in the early phases of experiencing GBS or CIDP by the longitudinally examined patients (36).

In general, acceptable validity, reliability, and responsiveness were demonstrated for SF-36 dimensions and summary measures in patients with immune-mediated polyneuropathies, hence emphasizing the clinical applicability of this scale as a generic health status in these conditions. To our knowledge, this study is the first that evaluated all psychometric properties of the SF-36 health status in immune-mediated polyneuropathies.

There are some methodological issues that should be addressed. First, because the longitudinally examined group of patients only consisted of 20 patients, it was decided not to analyse the patients with GBS and CIDP separately. Future studies are required to investigate the possible discriminatory validity of the SF-36 between these diagnostic categories and also between these categories and patients with MGUSP. However, we did not find any difference in quality of life using the SF-36 between patients with GBS versus those with CIDP or MGUSP who were stable for many years (35). Second, in the current study, participation of family members was stimulated in handling the patient's distress

caused by GBS or CIDP. A clear definition, however, of a patient's social network and the way family members should participate was not defined. Future studies in these conditions are essential to determine whether various forms of supportive social network (for example, family guidance alone versus family and psychological support) would lead to different outcomes. Despite these limitations, the comprehensiveness of the SF-36 generic health status may help to increase physicians' awareness, by providing information on functional health, general well-being, emotional state, and general health perceptions in patients with immune-mediated polyneuropathies.

References

1. Devinsky O. Outcome research in neurology. Incorporating health-related quality of life. Ann Neurol 1995;37:141-142.

2. Bernsen RAJAM, Jacobs HM, de Jager AEJ, *et al.* Residual health status after Guillain-Barré syndrome. J Neurol Neurosurg Psychiatry 1997;62:637-640.

3. Hadden RDM, Sharrack B, Bensa S, *et al.* Randomized trial of interferon β-1a in chronic inflammatory demyelinating polyradiculoneuropathy. Neurology 1999;53:57-61.

4. Molenaar DSM, Vermeulen M, de Visser M, *et al.* Impact of neurologic signs and symptoms on functional status in peripheral neuropathies. Neurology 1999;52:151-156.

5. Ruhland JL, Shields RK. The effects of a home exercise program on impairment and health-related quality of life in persons with chronic peripheral neuropathies. Phys Ther 1997;77:1026-1039.

6. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. J Neurol Neurosurg Psychiatry 1996;60:127-130.

7. Ware Jr. JE, Snow KK, Kosinski M, *et al.* SF-36 health survey. Manual and interpretation guide. Boston: The Health Institute, New England Medical Center, 1997.

8. Ware Jr. JE, Kosinski M, Keller SD. SF-36 physical and mental health summary scales: a user's manual. Boston: The Health Assessment Lab, New England Medical Center, 1994.

9. Aaronson NK, Muller M, Cohen PDA, *et al.* Translation, validation, and norming of the Dutch language version of the SF-36 health survey in community and chronic disease populations. J Clin Epidemiol 1998;51:1055-1068.

10. Ware Jr. JE, Gandek B, Kosinski M, *et al.* The equivalence of SF-36 summary health scores estimated using standard and country-specific algorithms in 10 countries: results form the IQOLA project. J Clin Epidemiol 1998;51:1167-1170.

11. Kleyweg RP, van der Meché FGA, Schmitz PIM. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré Syndrome. Muscle Nerve 1991;14:1103-1109.

12. Asbury AK and Cornblath DR. Assessment of current diagnostic criteria for Guillain- Barré Syndrome. Ann Neurol 1990;27(suppl):S21-S24.

13. Ad Hoc subcommittee of the American Academy of Neurology AIDS task force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

14. Miescher GC, Steck AJ. Paraproteinaemic neuropathies. Baillieres Clin Neurol 1996;5:219-232.

15. Merkies ISJ, Schmitz PIM, van der Meché FGA, *et al.* Psychometric evaluation of a new sensory scale in immune-mediated polyneuropathies. Neurology 2000;54:943-949.

16. Van der Meché FGA, Schmitz PIM, and the Dutch Guillain-Barré study group. A randomized trial comparing intravenous immune globulin and plasma exchange in Guillain-Barré syndrome. N Engl J Med 1992;326:1123-1129.

17. Hahn AF, Bolton CF, Zochodne D, *et al.* Intravenous immunoglobulin treatment in chronic inflammatory demyelinating polyneuropathy. A double-blind, placebo-controlled, cross-over study. Brain 1996;119:1067-1077.

18. Dyck PJ, Litchy WJ, Kratz KM, *et al.* A plasma exchange versus immune globulin infusion trial in chronic inflammatory demyelinating polyradiculoneuropathy. Annals Neurol 1994;36:838-845.

19. Van Doorn PA, Vermeulen M, Brand A, *et al.* Intravenous immunoglobulin treatment in patients with chronic inflammatory demyelinating polyneuropathy. Clinical and laboratory characteristics associated with improvement. Arch Neurol 1991;48:217-220.

20. Dyck PJ, O'Brien PC, Oviatt KF, *et al.* Prednisone improves chronic inflammatory demyelinating polyradiculoneuropathy more than no treatment. Ann Neurol 1982;11:136-141.

21. Nunnally JC. Psychometric theory. New York: McGraw Hill, 1978.

22. Herndon JE II, Harrell FE Jr. The restricted cubic spline hazard model. *Comm Stat Theory Methods* 1990;19:639-663.

23. Dwyer J, Feinleib M. Introduction to statistical models for longitudinal observation. In: Dwyer J, Feinleib M, Lippert P, Hoffmeister H, eds. Statistical models for longitudinal studies of health. New York, Oxford University Press, 1992: 3-48.

24. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;28:632-642.

25. Cohen J. *Statistical power analysis for the behavioural sciences*. New York: Academic Press; 1977.

26. Stoll TS, Gordon C, Seifert B, *et al.* Consistency and validity of patient administered assessment of quality of life by the MOS SF-36; its association with disease activity and damage in patients with systemic lupus erythematosus. J Rheumatol 1997;24:1608-14.

27. Peterson LS, Mason T, Nelson AM, *et al.* Psychosocial outcomes and health status of adults who have had juvenile rheumatoid arthritis. A controlled, population-based study. Arthritis Rheumatism 1997;40:2235-2240.

28. Deyo RA, Inui TS, Leininger J, *et al.* Physical and psychosocial function in rheumatoid arthritis. Arch Intern Med 1982;142:879-882.

29. Doering S, Henze T, Schussler G. Coping with illness in myasthenia gravis. Nervenartz 1993;64:640-647.

30. Kempen GIJ, Ormel J, Brilman EI, *et al.* Adaptive responses among Dutch elderly: The impact of eight chronic medical conditions on health-related quality of life. American J Public Health 1997;87:38-44.

31. Vichrey BG, Hays RD, Rausch R, *et al.* Quality of life of epilepsy surgery patients as compared with outpatients with hypertension, diabetes, heart disease, and/or depressive symptoms. Epilepsia 1994;35:597-607.

32. Hays RD, Cunningham WE, Sherbourne CD, *et al.* Health-related quality of life in patients with human immunodeficiency virus infection in the United States: results from the HIV Cost and Services Utilization Study. Am J Med 2000;108:714-722.

33. Eisendrath SJ, Matthay MA, Dunker JA, *et al.* Guillain-Barré syndrome: Psychosocial aspects of management. Psychosomatics 1983;24:465-475.

34. Krol B, Sanderman R, Suurmeijer TPBM. Social support, rheumatoid arthritis and quality of life: concepts, measurements and research. Patient Education and Counseling 1993;20:101-120.

35. Merkies ISJ, Schmitz PIM, Samijn JPA, *et al.* Fatigue in immune-mediated polyneuropathies. Neurology 1999;53:1648-1654.

36. Beusterien KM, Steinwald B, Ware Jr. JE. Usefulness of the SF-36 health survey in measuring health outcomes in the depressed elderly. J Geriatr Psychiatry Neurol 1996;9:13-21.

Chapter 11

Validity, reliability, and comparative responsiveness of impairment and disability measures in immune-mediated polyneuropathies

Merkies ISJ[1], Schmitz PIM[2], van Doorn PA[1], Samijn JPA[1], van der Meché FGA[1], for the Inflammatory Neuropathy Cause and Treatment (INCAT) group[3]

From the Department of [1]Neurology, [2]Department of Statistics - Daniel den Hoed Cancer Centre, Erasmus University Medical Centre Rotterdam, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands. [3]See Chapter 2 for INCAT-members.

Submitted for publication

Abstract

Objectives: To determine the validity, reliability, and rank ordering of the responsiveness of selected impairment and disability scales in sensory-motor immune-mediated polyneuropathies.
Methods: The validity and inter-/intra-observer reliability of three impairment scales (MRC sumscore, "*INCAT*" sensory sumscore, grip strength Vigorimeter) and three disability scales (Nine-hole peg test, Ten-metre walking test, an overall disability sumscore) were investigated in 113 patients with a stable neurological condition (83 patients who experienced Guillain-Barré syndrome (GBS) in the past, 22 with chronic inflammatory demyelinating polyneuropathy (CIDP), 8 with a polyneuropathy associated with a gammopathy of undetermined significance). For the assessment of responsiveness to clinical changes over time, all scales were utilized serially in twenty patients with recently diagnosed GBS (n = 7) or CIDP (n = 13) with changing clinical conditions. Responsiveness was measured using the: (1) Effect size; (2) Standardised response mean; (3) Wilcoxon matched pairs signed rank test; (4) newly devised Schmitz' distribution-free responsiveness score; (5) correlation studies between the scales' values and patients' rating of their clinical condition. For the responsiveness methods 1 to 4, the serially obtained values were plotted against a follow-up period of 52 weeks and the Area-under-the-curve (AUC) was calculated in each technique. In addition to literature suggestions regarding responsiveness values for the Effect size and Standardised response mean methods, good responsiveness was defined as an AUC > 41 (0.8 x 52 weeks) and moderate responsiveness as an AUC between 26 to 41 (0.5 to 0.8 x 52 weeks).
Results: Fair to good validity was demonstrated for all selected scales (in stable group: Spearman rank test: r = -0.19 to -0.63 and r = 0.48 to 0.69,$p < 0.05$; in longitudinally examined patients: linear regression analyses: R = -0.30 to -.089 and R = 0.27 to 0.81,$p < 0.0001$). Good inter- and intra-observer reliability was calculated for all scales (Anova: R = 0.85 – 0.99,$p < 0.0001$). In general, responsiveness techniques ranked the scales consistently. Responsiveness methods enabled the clinician to choose among equally valid and reliable outcome measures. The overall disability sumscore, MRC sumscore, and grip strength by the Vigorimeter were ranked among the best responsive scales and demonstrated good mutual association.
Conclusion: Good validity and reliability are provided for the selected impairment and disability measures. The use of the overall disability sumscore, MRC sumscore, and the Vigorimeter for grip strength assessment is primarily suggested for assessing outcome in patients with sensory-motor immune-mediated polyneuropathies, since these scales demonstrated the highest responsiveness scores.

Introduction

In 1980, the World Health Organization published a comprehensive framework, staging the consequences of any disease at various levels of outcome (1). This international classification of impairments, disabilities, and handicaps (ICIDH) allows evaluation and staging of clinical rating scales according to the level they represent (1,2). However, the value of outcome measures, not only depends on this categorization. Scales should also be appropriate to the patients group at study, brief and easily applicable, user friendly, communicable, and cost effective (2,3). Moreover, fulfilment of the clinimetric demands, like being valid, reliable, and responsive to clinical changes over time is required (3,4).
Validity reflects the relation between the concept to be measured and the scale used to assess that concept. Validity is usually established on expert judgements (content and face validity), by demonstrating a high correlation between the scale and a gold standard (criterion validity) or, in the absence of a gold standard, by examining the degree of association between a scale and other measures (3,4). Reliability addresses the internal consistency in multi-item scales and the ability of a scale to demonstrate reproducibility of the scores by the same (intra-) or a different examiner (inter-observer), or by the same patient (test-retest reliability) in case of self-rating scales (3,4). Responsiveness is defined as the ability of a scale to detect meaningful clinical changes over time when evaluating the benefits of a medical intervention (3,5,6). Responsiveness can be assessed within a group of patients receiving the same therapy or between groups of patients being treated with different therapy regimens (3,5,6).
Over the last decade, various easily applicable impairment and disability scales have been devised and used in clinical studies including patients with Guillain-Barré syndrome (GBS), chronic inflammatory demyelinating polyneuropathy (CIDP) or polyneuropathies associated with monoclonal gammopathies of undetermined significance (MGUSP). The most commonly use impairment scales are various motor scales based on the Medical Research Council (MRC) grading system and different sensory rating scales including various sensation modalities representing different sensory fibres (7). As disability scales, the Hughes' disability scale and related modifications, the (modified) Rankin scale, and the neurological disability score (NDS) have been used more often (8-12).
Surprisingly, despite the wide use of these scales, studies formally evaluating their clinimetric requirements are limited and incomplete. Regarding the motor scales, the MRC sumscore and the motor subset of the NDS are the only motor scales that have been validated and examined in terms of their reproducibility (9-11). With the exception of the sensory subset of the NDS, none of the used sensory scales has been subjected to a comprehensive clinimetric evaluation. The validity and reliability of the NDS sensory subset were demonstrated in diabetes patients with signs of a polyneuropathy (10,11). Good internal consistency was recently obtained for this scale in patients with hereditary and other polyneuropathies (13). The NDS sensory subset is, however, limited since sensory qualities are only assessed at the index finger and hallux.
The validity and reliability of the Hughes' disability scale were demonstrated in GBS patients (9). Despite its wide use in polyneuropathy studies, no formal clinimetric evaluation of the Rankin scale has been performed in patients with these disorders. Its reliability has been established in patients with stroke (12). Despite the simplicity and obvious face validity, the clinical use of the Hughes' scale and Rankin scale is limited, since their

emphasis is strongly directed towards mobility, thus not providing information regarding the arms.

Although validity and reliability have been demonstrated for some of the scales applied in patients with GBS, CIDP or MGUSP, relative little attention has been addressed towards their ability to detect clinically meaningful changes, an ability addressed as "responsiveness" (5,6). In addition, no consensus exists regarding which scale or set of scales reflects most appropriately the impairment and disability levels in patients with sensory-motor immune-mediated polyneuropathies.

Prompted by these findings, we evaluated the clinical applicability of existing scales that purely represented either the impairment or the disability outcome level in patients with immune-mediated polyneuropathies. This evaluation was performed by the "Inflammatory Neuropathy Cause and Treatment (*INCAT*) group", a collaborating force of European neurologists with special interest in immune-mediated polyneuropathies. The *INCAT* group evaluated various existing scales in terms of their content and face validity. A set of scales was selected for further studies. Subsequently, the construct validity, internal consistency in multi-items scales, inter- and intra-observer reliability for these measures were examined in patients with GBS, CIDP, and MGUSP (3,4). Also, because no information was available from the literature regarding which technique captures responsiveness the best of an outcome measure, we used various known responsiveness methods to examine this clinimetric requirement (5,6). The rank ordering by the responsiveness methods (from best to worst responsiveness) of the selected scales was investigated to determine whether responsiveness findings could help the clinician to select impairment en disability scales with the best responsiveness, among all scales that already demonstrated to be equally valid and reliable.

Patients and methods

Patients

One hundred and thirteen patients (83 GBS, 22 CIDP, 8 MGUSP) with a stable clinical condition were recruited from the Rotterdam immune-mediated polyneuropathy databank and the Dutch GBS study group (stable group). Patients with GBS, CIDP or MGUSP were recruited, since it was argued that these disorders represent parts of a continuum regarding their neuromuscular dysfunction pattern (14). The selected patients still had residual symptoms and/or signs due to their illness, representing a broad range of disability. Nine CIDP patients required interval treatment ranging from weeks to months, with intravenous immunoglobulin (IVIg). With this therapy their clinical condition has been stable for more than 6 months. Six patients with MGUSP (three with IgG, two with IgM, and one patient with IgG+IgM) had an associated demyelinating polyneuropathy with minor concurrent axonal damage in three. An axonal polyneuropathy was diagnosed in the remaining two patients with MGUSP (one IgA and one IgG type).

Twenty consecutive patients with recently diagnosed sensory-motor GBS (n = 7) or CIDP (n = 13) and changing clinical conditions were enrolled to investigate the responsiveness of the selected scales (longitudinal group). All GBS and CIDP patients met the international criteria for their illness (15,16). The diagnosis MGUSP was established after excluding all possible causes for the gammopathy and polyneuropathy (17).

Literature review and scales selection

From January 1988 till January 1999 a systematic Medline search was performed reviewing all impairment and disability methods in clinical neuromuscular studies. A set of twelve scales was selected representing the impairment and disability levels of outcome, taking into account the clinical spectrum of patients with GBS, CIDP, and MGUSP (14). This set of scales was presented to and evaluated by an expert panel consisting of 13 neurologists, all *INCAT* members, who selected three impairment and three disability scales for further clinimetric evaluation in sensory-motor immune-mediated polyneuropathies. Since, strength and sensation are primarily affected in these conditions, the MRC sumscore and "*INCAT*" sensory sumscore were chosen to capture a global view of these impairment entities (9,18). In addition, distal weakness may predominate in these disorders and therefore the easily applicable hand-held Vigorimeter (VM) was used to measure grip strength (19). Also, grip strength has been demonstrated to be a prognostic indicator of clinical and functional recovery and useful in monitoring the effect of treatment (9,20-23). To assess "focal" disability, the Nine-hole peg (dexterity) test and the Ten-metre walking (mobility) test were chosen (2). "Global" (arm + leg) disability was measured using an overall disability sumscore (24).

Assessment tools/scales

The *MRC Sumscore* is a summation of the MRC grades (range: 0 - 5) given in full numbers of the following muscle pairs: upper arm abductors, elbow flexors, wrist extensors, hip flexors, knee extensors, and foot dorsal flexors (9). The MRC sumscore ranges from 0 ("total paralysis") to 60 ("normal strength") (9). Good validity and inter-observer reliability are provided for this scale in patients with GBS (9).

The *"INCAT" sensory sumscore* (ISS) was recently introduced and extensively evaluated in patients with immune-mediated polyneuropathies (18). In brief, the ISS comprises pinprick and vibration sense plus a two-point discrimination value in the arms and legs and ranges from 0 ("normal sensation") to 20 ("most severe sensory deficit") (18).

The *Vigorimeter* (VM) (Martin, Tuttlingen, Germany) is a hand-held dynamometer to measure grip strength (19). The pressure in the bulb is registered on a manometer via a rubber junction tube and expressed in kiloPascals (kPa). The medium sized bulb was applied in the selected participants. Good validity for the VM was reported after correlation with the Jamar dynamometer (25,26). Good reliability and responsiveness values were reported for this device in healthy individuals and patients with rheumatoid arthritis (27,28).

The *Nine-hole peg* (dexterity) test and the *Ten-metre walking* (ambulatory) test were assessed to measure "focal" disability (2). The simplicity, validity, and reliability of these tests have been demonstrated, particularly in patients with stroke (2).

The *Overall disability sumscore* (ODSS) is composed by a recently published arm and leg disability scales with a total score ranging from 0 ("no signs of disability") to 12 ("most severe disability score") (24). It comprises a good functional description of the arms and legs in a checklist form suitable for interviewing patients. Daily arm activities like dressing upper part of the body, doing and undoing buttons and zips, washing and brushing hair, using a knife and fork and turning a key in a lock are scored as being "not affected", "affected but not prevented" or "prevented". Subsequently, these results are translated into an arm grade (score range: 0 [normal arm abilities] to 5 [severe symptoms and signs in both arms preventing all purposeful movements]). The leg scale highlights problems regarding

walking taking into account the use of a device. The results are also translated into a leg grade (score range: 0 [walking is not affected] to 7 [restricted to wheelchair or bed most of the day, preventing all purposeful movements of the legs]) (24). The selected arm and leg disability scales are subsets of a more comprehensive Guy's neurological disability scale (24). Good clinimetric requirements have been recently demonstrated for all components of the Guy's scale in patients with multiple sclerosis (24).

Test procedures

General aspects. All participants gave informed consent prior to the study. All measures were obtained in a quiet, comfortably warm, and temperature-controlled room (approximately 20°C) at our outpatient clinic. The assessments were performed in a random order. For the assessment of strength, a standardised joint and limb position as well as the point at which counter-force was administered was defined before the start of the study and taken at examination of each muscle group (page 205). Sensory modalities were examined in triplicate under the earlier prescribed standard conditions with the patients lying in supine position (18).

For the assessment of grip strength with the VM, all patients were examined according to the assessment recommendations by the American Society of Hand Therapists (29). In brief, the participants were seated with their shoulder adducted and neutrally rotated, elbow flexed at 90°, forearm in neutral position, and wrist between 0° and 30° dorsiflexion and between 0° and 15° of ulnar deviation (29). Three grip strength measurements with maximum voluntary contractions for each hand were taken in alternating order. Between each trial a pause of 30 seconds was assigned. The results of three trials for each hand were averaged and considered the grip strength score for that particular hand.

All patients received training in assessing the Nine-hole peg test prior to the start of the study in order to exclude any training effect. The patients were asked, under the prescribed standard conditions, to pick up nine pegs from a tray at table height and place them as quickly as possible into nine holes in a neighbouring horizontal board. After this procedure, the pegs were removed as fast as possible. These measures were performed for both hands separately in alternating order and the time required to fulfil these tasks was recorded (in seconds) (2). Patients were also requested to walk ten meters in a straight line at their preferential speed, using whatever aid-needed (2). Three measures were completed for each of these tests and the corresponding time was recorded at each assessment (in seconds). For each of these two tests, the mean time of completion was calculated by averaging the three obtained measures. The study took place between January 1999 and January 2000.

Validity and reliability. The first step in the assessment of validity was an extensive evaluation by the "*INCAT*" expert panel of reported impairment and disability outcome measures that were applied in immune-mediated polyneuropathy studies. Eventually, this panel selected the above-presented scales and instruments. The selection was based on the presumed conjoint ability of this set of scales to cover the most important clinical aspects of patients with GBS, CIDP, and MGUSP. Construct validity of the selected scales was investigated by correlation and regression studies.

For the assessment of reliability and construct validity of the selected scales in the stable group of 113 patients, two neurologists and six experienced residents in neurology formed 28 different couples. Preceding the study, all investigators received instructions in assessing the outcome measures. Twenty-seven ("variable") couples investigated a total of 68 patients

(2-3 patients for each couple). The remaining 45 stable patients were investigated by the "experienced" couple (IM+ JS). The latter couple was formed to examine the effect of training (and thus a possible increase in reliability) when using the scales often.
The patients were examined on two different occasions at our outpatient clinic. During the first visit the two members of an appointed pair performed their scores independently and consecutively (usually within 2 hours) (inter-observer measures). Within 2-4 weeks, the patient returned for a second visit and only one investigator of the earlier assigned pair examined the patient again (intra-observer values) without having access to previous results. The assessments sequence at entry and the examination at the second visit were equally distributed among the members of an assigned couple. Eventually, each member of a couple examined the approximately same number of patients. All scales were assessed at each visit in all patients.
Responsiveness. Twenty consecutive patients were longitudinally examined by the same clinician (ISJM) and all scales were assessed at study entry and 8-13 times during follow-up. There was a standard follow-up schedule (week 0, 2, 4, 8, 12, 16, 21, 26, 32, 40 and 52) with additional clinical investigations if necessary. At each visit, the patients were requested to judge whether their clinical condition deteriorated (grade 1), remained stable (grade 2) or improved (grade 3) when compared with the last visit ("clinical-judgement scores"). At study entry, the patients reflected their clinical condition against their physical status within the two weeks before the start of the study.

Statistics
Validity and Reliability. In the stable group of patients, the correlation between the scales was analysed using the Spearman rank correlation test. In the longitudinal group, random effects linear regression analyses between the scales were performed taking into account the correlation of the data caused by the longitudinal structure. The latter was achieved using the program "xtreg" in STATA 6.0 for Windows 95 which is based upon a cross-sectional time-series regression model as described by Dwyer and Feinleib (StataCorp. 1997. Stata Statistical Software: Release 5.0. College Station, TX: Stata Corporation; see also reference 30). A logarithmic transformation was applied to various scale values prior to the regression studies to strive for the best fit. This was achieved by systematic evaluation of the possible correlation between the values of two scales through constructed graphs. Cronbach's alpha coefficient was estimated for the MRC sumscore, sensory sumscore, and overall disability sumscore in both (stable and longitudinal) patients groups (31). The interobserver and intraobserver reliability for the scales was quantified by estimation of the intraclass correlation coefficient using a one-way random effects analysis-of-variance model for the two investigator ("experienced" and "variable") groups.
Responsiveness. The selected scales' values could not be always collected at the exact follow-up date, due to practical patients' and investigators' inconveniences. Therefore, where necessary, interpolation of the data was performed in each patient. Subsequently, corresponding values for each scale in each patient were calculated for the exact weeks 0, 2, 4, to up to 52. Responsiveness was than evaluated in five ways: (1) effect size, (2) standardised response mean, (3) Wilcoxon matched pairs signed rank test, (4) Schmitz' distribution-free responsiveness score, and (5) correlation studies between the scales' values and patients' rating of their clinical condition ("clinical-judgement scores"). The effect size is the mean change in scores divided by the standard deviation of the baseline scores (effect

size = (μi - μo)/SDμo; μi = mean scale value of the longitudinally examined group at week = i; μo = mean scale value at week = 0 [entry]; see also reference 32). The standardised response mean is equal to the mean change in scores divided by the standard deviation of change in scores (standardised response mean = (μi - μo)/SD(μi - μo); μi = mean scale value of the longitudinally examined group at week = i; μo = mean scale value at week = 0 [entry]; see also reference 33). The Wilcoxon matched pairs signed rank test is the equivalent distribution-free method of the relative efficiency test postulated by Liang and associates (34). The Schmitz' distribution-free responsiveness score is defined as 1.35 x the median change in scores divided by the inter-quartile range of change in scores (IQRchange = 75th percentile - 25th percentile). The last two methods were new and incorporated, since all available responsiveness methods thus far are based upon parametric statistics. Finally, change in scores was correlated with an external criterion, the clinical-judgement scores using "xtreg" (intraclass correlation coefficient [R] are presented) (30).

The obtained responsiveness values for the methods (1) to (4) were serially plotted against time. Subsequently, the area under the curve (AUC) of each graph was calculated, making it possible to compare the AUC's for the selected scales in each responsiveness technique and therefore obtain a rank ordering. Moreover, these graphs provided information whether the longitudinally obtained responsiveness values had a more 'static' or 'modulating' pattern. A bigger AUC was defined as a higher responsiveness value for the corresponding scale. Moreover, according to Cohen, an effect size or standardised response mean value between 0.5 and 0.8 is considered moderate, and 0.8 or greater is considered as high responsiveness (35). Based on these premises, a moderate area-responsiveness was calculated for these two methods by multiplying 0.5-0.8 with 52 weeks of follow-up. Moderate area-responsiveness corresponded therefore with an AUC between 26 - 41. Good area-responsiveness for the effect size and standardised response mean was defined as AUC > 41 (calculated by multiplying 0.8 with 52 weeks of follow-up).

For each scale, median values at 12, 26, 40, and 52 weeks of follow-up were compared with the median value at entry (Wilcoxon signed-rank test) to demonstrate possible changes in time. All analyses were performed using Stata 6.0 for Windows 95. A value of $p \leq 0.05$ was considered statistically significant.

Results

General aspects, Content and Face Validity. The selected scales were thoroughly evaluated by a panel of 13 neurologists who eventually concluded that they all have content and face validity. Subsequently, the construct validity, reliability, and responsiveness for these scales were assessed.

All eight examiners who investigated the patients concluded that the selected scales were easy in use. The total examination of a patient took twenty to thirty minutes to be completed.

The stable group of patients (54 females; 59 males; median age 56, range 14-84 years) had a median duration of symptoms till onset of the study of 5.1 years. Seven of these patients were bed bound and fourteen patients required assistance or a device to walk short distances. The remaining 92 patients could walk independently. The corresponding median values and

Table 1

Median values (range) of impairment and disability measures in the stable group of 113 patients with sensory-motor immune-mediated polyneuropathies

	First assessment	Second assessment	Third assessment
MRC-sumscore	56 (16 – 60)	54 (18 - 60)	56 (16 - 60)
"*INCAT*" sensory sumscore	3 (0 – 15)	3 (0 - 18)	3 (0 - 18)
GS-RH (kPa)	67 (0 – 156)	65 (0 - 158)	67 (0 - 152)
GS-LH (kPa)	64 (0 – 158)	62 (0 - 160)	62 (0 - 158)
9HPT-RH (in seconds) *	24.4 (15.4 - 134.7)	23.1 (14.6 - 143.6)	23.7 (14.2 - 144.8)
9HPT-LH (in seconds)*	23.8 (15.8 - 192.2)	24.5 (17.1 – 196.2)	23.6 (15.2 - 144.6)
10MWT (in seconds)**	8.3 (5.4 – 32.0)	8.2 (5.4 – 34.0)	8.3 (4.6 - 34.2)
Disability-ss	4 (0 – 11)	4 (0 – 12)	3 (0 - 12)

Legend to Table 1
GS = grip strength; 9HPT = Nine-hole peg test; 10MWT = 10 meters walking test; RH = Right hands; LH = Left hands. *Five patients could not fulfil the Nine-hole peg test. **Seven patients were not able to walk. The MRC sumscore ranges from 0 ("total paralysis") to 60 ("normal strength") (9). The "*INCAT*" sensory sumscore ranges from 0 ("normal sensation") to 20 ("most severe sensory deficit") (18). The Vigorimeter values range from 0 (lowest grip strength) to 160 kiloPascals (kPa) (highest grip strength) (19). The overall disability sumscore ranges from 0 ("no signs of disability") to 12 ("most severe disability score") (24).

ranges for all scales in these patients are presented in table 1. All measures except grip strength values demonstrated a non-Gaussian distribution.

Construct Validity, Internal consistency, inter- and intra-observer Reliability. The correlation studies between the selected scales in the stable group and the regression analyses performed between these measures in the longitudinal group are presented in table 2. In general, moderate to good correlation values were obtained for each scale, thus demonstrating the validity of all selected scales. In the stable group, the strongest correlation was obtained between the Ten-metre walking test and the overall disability sumscore ($r = 0.69$; $p < 0.0001$). In the longitudinally examined patients the highest association was demonstrated between the MRC sumscore and the overall disability sumscore ($R = 0.89$; $p < 0.0001$).

In the stable group, acceptable internal consistency values were obtained for the multi-item scales (MRC sumscore: 0.94 at first and third assessments, 0.93 at second assessment; "*INCAT*" sensory sumscore: 0.68, 0.73, and 0.71 at first, second, and third assessments, respectively; overall disability sumscore: 0.72, 0.70, and 0.76 at first, second, and third assessments, respectively). The internal consistency in the longitudinally examined patients was 0.96, 0.86, and 0.78 for the MRC sumscore, "*INCAT*" sensory sumscore, and overall disability sumscore, respectively. As shown in table 3, good inter- and intra-observer reliability values were obtained for the scales by the "experienced" and "variable" couples of investigators (intraclass correlation coefficients ranging from $R = 0.85$ to 0.99).

Responsiveness. Eight females and twelve males (median age 54.0, range 15 - 70 years) were examined longitudinally. At study entry, four patients were bed bound, one requiring artificial ventilation, and nine patients were unable to walk independently. All patients experienced general loss of strength, sensory disturbances, and deficit in daily functional activities. Two hundred and one visits were completed during a follow-up period of 40 to 58 (median: 52) weeks. Nineteen patients completed a one-year follow-up. With the exception

Table 2

Correlation and regression analyses between impairment and disability measures in patients with sensory-motor immune-mediated polyneuropathies

	MRC-ss	Sensory-ss	GS-RH	GS-LH	9HPT-RH	9HPT-LH	10MWT
Stable group of patients (n=113)							
Sensory-ss	-0.19**						
GS-RH	0.49*	-0.35*					
GS-LH	0.56*	-0.37*	-				
9HPT-RH$^{\lambda}$	-0.34*	0.52*	-0.46*	-0.45*			
9HPT-LH$^{\lambda}$	-0.35*	0.49*	-0.43*	-0.45*	-		
10MWT$^{\pi}$	-0.38*	0.48*	-0.51*	-0.56*	0.61*	0.57*	
Disability-ss	-0.59*	0.50*	-0.62*	-0.63*	0.59*	0.61*	0.69*
Longitudinal group of patients (n=20; 201 visits)							
Sensory-ss	-0.30*						
GS-RH	0.76*	-0.46*					
GS-LH	0.73*	-0.48*	-				
9HPT-RH$^{\infty}$	-0.53*	0.50*	-0.63*	-0.54*			
9HPT-LH$^{\infty}$	-0.58*	0.63*	-0.64*	-0.59*	-		
10MWT$^{\Phi}$	-0.63*	0.41*	-0.32*	-0.28*	0.27*	0.36*	
Disability-ss	-0.89*	0.74*	-0.72*	-0.69*	0.75*	0.81*	0.66*

Legend to Table 2

GS = grip strength; 9HPT = Nine-hole peg test; 10MWT = 10 meters walking test; RH = Right hands; LH = Left hands; ss = sumscore. **p < 0.05; * p < 0.001.

In the stable group: The presented data represent the obtained Spearman rank correlations (r) at entry between scales' values. These correlations were almost identical at second and third assessments and therefore only once presented. $^{\lambda}$Five patients could not fulfil the Nine-hole peg test. $^{\pi}$Seven patients were not able to walk.

In the longitudinal group: The presented data represent the obtained associations between the scales' values using the program "xtreg" (see also section Statistics; 30). The values represent intraclass correlation coefficients (R). $^{\infty}$In 2 patients, at six occasions the Nine-hole peg test could not be performed due to weakness and sensory disturbances. $^{\Phi}$Five patients were not able to walk at a total of ten measurements. Note: correlations may be negative or positive.

of one GBS patient who only experienced mild symptoms, all patients received initial treatment with IVIg (0.4 grams/kilogram body weight/day for 5 consecutive days). All but one patient with CIDP showed good functional improvement on IVIg during follow-up. The non-responder received a treatment course of oral prednisone, 100 milligrams/day for four consecutive weeks. This patient improved also with this therapy and prednisone was tapered down over five months period to 30 milligrams on alternate days.

The GBS patients did not show any deterioration and improved gradually during follow-up. After initial improvement, all 12 IVIg responsive CIDP patients showed deterioration in

their clinical condition. Consequently, maintenance therapy with IVIg (1-2 days 0.4 grams/kg/day; intervals: 3 - 21 weeks) was needed to prevent further deterioration and to regain earlier achieved improvement. Eventually, all patients demonstrated during follow-up a general decrement in impairment and improvement of functional abilities. Improvement was detected on all scales as presented in table 4.
The patients graded their clinical condition 53 times as "deteriorating", 38 times as "stable" and 110 times as "improving". Figures 1 and 2 correspond with the standardised response mean scores and Schmitz' distribution-free responsiveness scores in time. The AUC's in these two figures demonstrated that the overall disability sumscore had the highest responsiveness values followed by the MRC sumscore and Grip strength values obtained with the Vigorimeter. Also, as can be seen, the serially obtained area-responsiveness values demonstrated a modulating pattern over time. The calculated AUC's for the scales in each responsiveness method and the random effects linear regression studies between the scales' values and the clinical-judgement scores are presented in table 5. In general, the responsiveness indices ranked the scales consistently. The overall disability sumscore turned out to be the best responsive scale. Also, The MRC sumscore, the Vigorimeter, and the "*INCAT*" sensory sumscore demonstrated also acceptable responsiveness, reaching effect size and standardised responsiveness mean AUC's ≥ 39. The worst responsiveness scores were obtained in the Nine-hole peg test and Ten-metre walking test (see Figures 1 and 2 and Table 5). The Ten-metre walking test demonstrated the worst responsiveness area in the Schmitz' distribution-free method.

Table 3

Reliability of impairment and disability scales in a stable group of patients with sensory-motor immune-mediated polyneuropathies

	"Experienced" couple of examiners (couple number 1) 45 patients				"Variable" couples of examiners (couples number 2-28) 68 patients			
	Intraclass correlation coefficient **R** + (SD) ($p < 0.0001$ for all associations)				Intraclass correlation coefficient **R** + (SD) ($p < 0.0001$ for all associations)			
Impairment scales	**Interobserver**		**Intraobserver**		**Interobserver**		**Intraobserver**	
MRC sumscore	0.93	(0.02)	0.93	(0.02)	0.86	(0.02)	0.95	(0.008)
"*INCAT*" sensory sumscore	0.89	(0.02)	0.85	(0.03)	0.86	(0.02)	0.87	(0.02)
Grip strength (Right hands)	0.97	(0.007)	0.96	(0.008)	0.97	(0.005)	0.97	(0.006)
Grip strength (Left hands)	0.96	(0.009)	0.95	(0.01)	0.97	(0.005)	0.96	(0.006)
Disability scales								
Nine-hole peg test (Right hands)	0.93	(0.02)	0.95	(0.01)	0.96	(0.006)	0.90	(0.02)
Nine-hole peg test (Left hands)	0.93	(0.02)	0.89	(0.02)	0.97	(0.004)	0.94	(0.01)
Ten-metre walking test	0.99	(0.02)	0.85	(0.03)	0.93	(0.01)	0.96	(0.007)
Overall disability sumscore	0.95	(0.01)	0.95	(0.01)	0.90	(0.02)	0.93	(0.01)

Discussion

The current study was designed to evaluate a set of impairment and disability scales in order to provide a more standardised approach in choosing endpoints for clinical trials in sensory-motor immune-mediated polyneuropathies. The evaluation was focused not only on validity and reliability, but also reflected the evaluation of the selected scales' responsiveness to

clinical changes over time, since of all clinimetric requirements this item has been the least studied in the evaluation of outcome measures (5).

All selected scales demonstrated acceptable construct validity and reliability, thus reflecting the good judgement (content and face validity) by the experts' opinion. Whereas validity and reliability form the clinimetric core stones of a rating scale, the ability of a measure to detect clinically meaningful changes over time is crucial (2,5,6,36). For clinicians and researchers, such a measure should discriminate between irrelevant changes (normal fluctuations in the activity of an illness; "noise") and clinically meaningful changes on which a treatment policy can be based ("signal") (5,6). A statistic and heuristic approach have been proposed by Liang to examine responsiveness of outcome measures (5). Statistical responsiveness captures the ability of an instrument to measure any change. Methods 1 to 4 applied in the current study are examples of statistical responsiveness techniques. Heuristic responsiveness techniques relate changes as assessed by an outcome measure to an external indicator (e.g. the clinical-judgement scores by the patients in the current study; see also references 5 and 6). We examined these two approaches using various statistical responsiveness techniques and regression analyses studies between the selected impairment and disability scales' values and the clinical-judgement scores (Table 5). Responsiveness enabled us to differentiate between equally valid and reliable scales and to compare the results of the various methods applied, particularly between the parametric and non-parametric responsiveness techniques. The non-parametric methods (Schmitz' distribution-free responsiveness scores or Wilcoxon matched pairs signed rank test) applied in the current study are believed to capture "responsiveness" the best, because all scales except grip strength by the Vigorimeter demonstrated a non-Gausian distribution pattern (Table 1).

Table 4

Median (95% inter-percentile range) values of impairment and disability scales in longitudinally examined patients with sensory-motor immune-mediated polyneuropathies

	At entry	At week 12	At week 26	At week 40	At week 52
MRC-ss	50.5 (14 – 58)	58 (28 - 60)*	59 (30 - 60)*	59.5 (32 – 60)*	60 (37 - 60)*
"*INCAT*" sensory-ss	7.5 (1 - 18)	3 (0 – 13.5)*	2 (0 - 12)*	3 (0 - 18)*	2 (0 – 16)**
GS-RH	54 (0 - 103)	73 (8 - 121)*	81 (12 - 159)*	90 (8 – 160)*	82 (0 – 160)**
GS-LH	53 (0 - 94)	75 (6 - 120)*	79 (9 - 158)*	88 (6 – 160)*	87 (0 – 160)**
9HPT-RH [~]	25.8 (19.5 – 76.3)	23.8 (15.0 - 96.8)**	18.3 (15.3 - 38.2)*	19.2 (14.4 – 141.9)*	20.3 (15.4 – 146)**
9HPT-LH [~]	29.2 (18.2 – 60.5)	26.7 (17.4 – 88.2)*	20.9 (16.7 - 42.9)*	20.5 (15.6 – 109)**	22.1 (15.6 – 145)[γ]
10MWT[◊]	10.0 (6.3 - 32.0)	7.4 (5.3 - 13.4)*	6.9 (5.0 - 15.1)*	6.8 (5.7 - 11.0)*	7.1 (5.1 – 11.3)**
Disability-ss	5 (3 - 11)	3 (0 - 10.5)*	2 (0 – 9.5)*	2 (0 - 9)*	2 (0 – 9)**

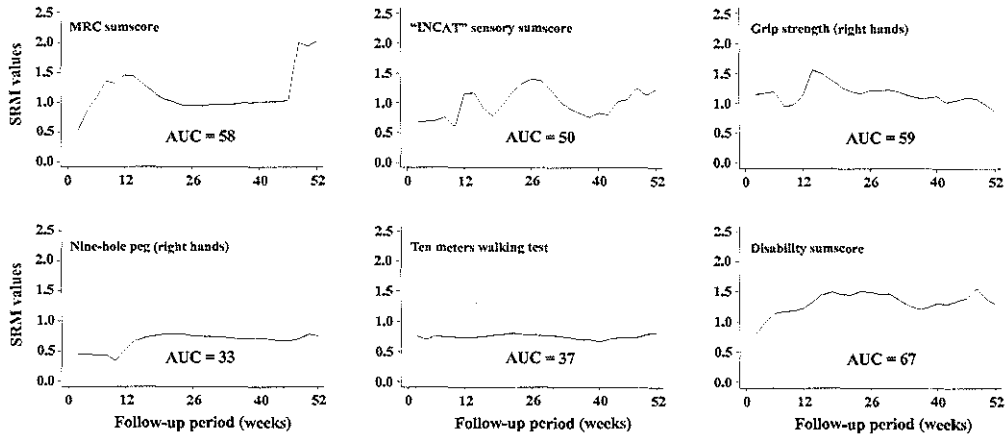
Legend to Table 4

GS = grip strength; 9HPT = Nine-hole peg test; 10MWT = 10 meters walking test; RH = Right hands; LH = Left hands; ss = sumscore. Twenty patients were examined longitudinally. Two hundred and one visits were completed. [~]In 2 patients, at six occasions the Nine-hole peg test could not be performed due to weakness and sensory disturbances. [◊]Five patients were not able to walk at a total of ten measurements. For each scale, median values at 12, 26, 40, 52 weeks of follow-up were compared with median value at entry (Wilcoxon signed-rank test). **p ≤ 0.01; *p ≤ 0.001; [γ] = not significant. Improvement on the MRC sumscore and grip strength was characterized by an increase in scores.

A reduction in the scores for the "*INCAT*" sensory sumscore, Nine-hole peg test, Ten-metre walking test, and overall disability sumscore reflected improvement.

Figure 1

Standardised response mean (SRM) values with corresponding area-under-the-curve (AUC) for impairment and disability scales in patients with sensory-motor immune-mediated polyneuropathies



Legend to Figure 1
Moderate area-responsiveness was defined as an AUC between 26 – 41 and good area-responsiveness as AUC > 41 (see section statistics). A bigger AUC corresponds with higher area-responsiveness.

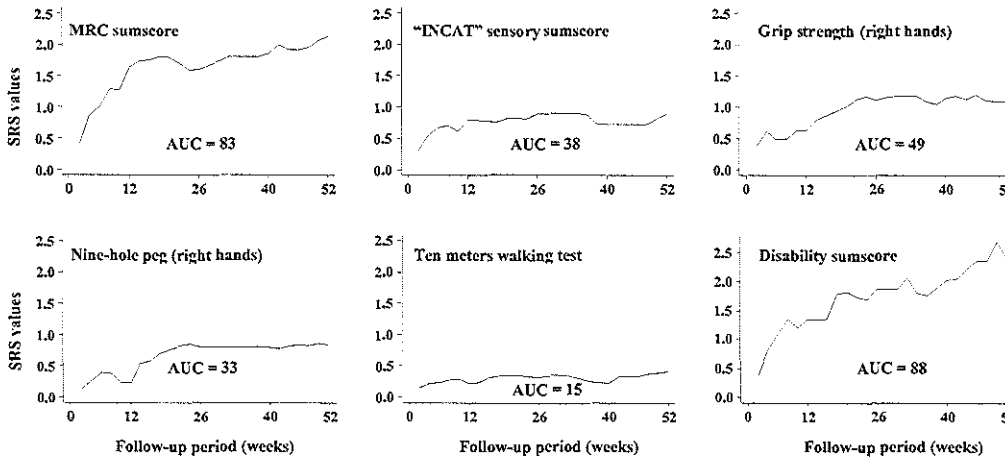
Hence, complementary to literature findings, these non-parametric methods are more robust when evaluating responsiveness. The use of the overall disability sumscore, the MRC sumscore, and the Vigorimeter is primarily recommended in the follow-up of patients with sensory-motor immune-mediated polyneuropathies, because these scales demonstrated to be the most responsive ones.

As far as we know, this is the first neurological paper addressing the more dynamic responsiveness technique of Area-under-the curve for outcome measures. The validity of this approach however has been demonstrated in rheumatoid arthritis (37). In our view, this approach has extended the postulated definitions of moderate and good responsiveness by Cohen for the effect size and standardised response mean methods (35). Also, traditional papers addressing scales' responsiveness generally report the responsiveness values at one or two arbitrarily chosen post-medical intervention moments (32-34,38-40). Thus, no longitudinal reflection of the "true" responsiveness over time of these measures is provided. In contrast with the more "static" approach, the graphical representation in the current study of the serially obtained responsiveness values demonstrates that responsiveness is a modulating dynamic entity (Figures 1 and 2).

In the current study, the Ten-metre walking test and the Nine-hole peg test were both easily applicable with good validity and highest reliability values. However, the responsiveness of these two scales was very poor (Tables 2, 3, 5 and Figures 1 and 2). This implicates that good validity and reliability of an outcome measure alone will not suffice to ensure relevant clinical applicability. The consequence of a less responsive instrument is that the number of patients that is required to achieve a given statistical power becomes higher (36).

Figure 2

Schmitz's distribution-free responsiveness (SRS) values with corresponding area-under-the-curve (AUC) for impairment and disability scales in patients with sensory-motor immune-mediated polyneuropathies



Legend to Figure 2
A bigger AUC corresponds with a higher area-responsiveness.

Table 5

Area-under-the-curve (AUC) values and regression analyses representing responsiveness of impairment and disability scales in patients with sensory-motor immune-mediated polyneuropathies (n=20)

	Effect size	Standardised response mean	Wilcoxon matched pairs signed rank test	Schmitz' distribution-free responsiveness score	Random effects linear regressions with clinical-judgement Scores (R) "xtreg"
	Mean change / SDbaseline	Mean change / SDchange		1.35 x Median change / IQRchange	
MRC-ss	39	58	188	83	0.46
"*INCAT*" sensory-ss	41	50	157	38	0.51
GS-RH	55	59	180	49	0.50
GS-LH	47	55	177	46	0.46
9HPT-RH	27	33	151	33	0.24
9HPT-LH	37	44	155	39	0.23
10MWT	35	37	160	15	0.28
Disability-ss	61	67	188	88	0.46

Legend to Table 5
GS = grip strength; 9HPT = Nine-hole peg test; 10MWT = 10 meters walking test; RH = Right hands; LH = Left hands; ss = sumscore. IQR = inter-quartile range (75th percentile minus 25th percentile). For the effects size and standardised response mean methods, moderate area-responsiveness was defined as an AUC between 26 – 41 and good area-responsiveness as: AUC > 41 (see section statistics). Random effects linear regression analyses between the scales' values and clinical-judgement scores were performed, taking into account the correlation of the data caused by the longitudinal structure using a serial regression method ("xtreg"; see section statistics; 30).

With respect to the aims of the current study, some methodological issues should be addressed. First, despite a large amount of visits in the longitudinally examined patients (a total of 201), this group consisted of only twenty patients. Some caution is required in interpreting the responsiveness values, since it is not determined whether a larger group of patients would give the same rank order of the selected scales. Second, not all available impairment and disability scales have been incorporated in the current study. However, based on the experts' judgement, it is believed that the selected scales covered the main clinical aspects of patients with GBS, CIDP, and MGUSP taking into account their simplicity and communicability. Also, because patients with MGUSP have a more indolent course of their disease compared with patients with GBS or CIDP, it is conceivable that the selected scales will not demonstrate the same responsiveness values over a short period of time. Therefore, some caution is required in extrapolating the use of the evaluated measures in the current study to patients with MGUSP. Third, the obtained responsiveness scores for the methods (1) to (4) only demonstrated within-group responsiveness for the selected scales. It is not clear whether substantial discriminative responsiveness scores will be obtained for these scales when evaluating various groups of patients, for example, in a trial setting comparing a placebo versus a treated group (36). Also, it should be stated that a statistically significant difference between treated and control patients does not necessarily mean a clinically significant difference. Future studies are required in patients with sensory-motor immune-mediated polyneuropathies to determine the minimal clinical important difference for the selected outcome measures (41). It has to be determined as well whether this minimal clinical important difference should be based on experts' opinion, within-patients' or between-patients' judgments (41). Fourth, a scale with the greatest statistical responsiveness may not always be measuring changes that are most important to patients (39). For example, the overall disability sumscore had the highest statistical responsiveness, but did not show the highest association with the patients' clinical-judgement scores. Instead, the "*INCAT*" sensory sumscore demonstrated the highest heuristic responsiveness with lower statistical responsiveness values as compared with the overall disability sumscore. The question remains open whether more attention should be focused on the results related to the patients' perspective of their illness despite a lower statistical responsiveness. Moreover, not only an optimal fulfilment of all clinimetric demands will function as the only passport to include a scale in a study. There may be other reasons as well to incorporate a scale in a study. For example, a scale can be included in a study if one wants to gather information regarding a specific quality (e.g. sensory deficit changes over time related to therapy).

In conclusion, good validity and reliability are demonstrated for a set of impairment and disability measures in patients with sensory-motor immune-mediated polyneuropathies. Based upon statistical and heuristic responsiveness methods, a general consistent rank ordering of the evaluated outcome measures was demonstrated in these conditions, thus enabling the clinician to choose among equally valid and reliable scales. The use of the overall disability sumscore, MRC sumscore, and grip strength by the Vigorimeter is suggested for assessing outcome in patients with sensory-motor immune-mediated polyneuropathies, since these three outcome measures showed the highest responsiveness values.

References

1. World Health Organization. International classification of impairments, disabilities, and handicaps. Geneva: WHO, 1980.

2. Wade DT. Measurements in neurological rehabilitation. Oxford University Press: Oxford, 1992.

3. Hobart JC, Lamping DL, Thompson AJ. Evaluating neurological outcome measures: the bare essentials. J Neurol Neurosurg Psychiatry 1996;60:127-130.

4. Feinstein AR. Clinimetrics. Yale University Press: New Haven and London, 1987.

5. Liang MH. Evaluating measurement responsiveness. J Rheumatol 1995;22:1191-1192.

6. Fortin PR, Stucki G, Katz JN. Measuring relevant change: an emerging challenge in rheumatologic clinical trials. Arthritis Rheum 1995;38:1027-1030.

7. Medical Research Council. Aids to the investigation of the peripheral nervous system. Her Majesty's Stationary Office, London. 1943;1-2.

8. Hughes RAC, Newsom-Davis JM, Perkin GD, *et al.* Controlled trial prednisolone in acute polyneuropathy. Lancet 1978;2:750-753.

9. Kleyweg RP, van der Meché FGA, Schmitz PIM. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré syndrome. Muscle Nerve 1991;14:1103-1109.

10. Dyck PJ, Kratz KM, Lehman KA, *et al.* The Rochester diabetic neuropathy study: Design, criteria for types of neuropathy, selection bias, and reproducibility of neuropathic tests. Neurology 1991;41:799-807.

11. Dyck PJ, Karnes JL, O'Brien PC, *et al.* The Rochester diabetic neuropathy study: Reassessment of tests and criteria for diagnosis and staged severity. Neurology 1992;42:1164-1170.

12. van Swieten JC, Koudstaal PJ, Visser MC, *et al.* Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988;19:604-607.

13. Molenaar DS, Vermeulen M, de Visser M, *et al.* Impact of neurologic signs and symptoms on functional status in peripheral neuropathies. Neurology 1999;52:151-156.

14. Hartung H-P, van der Meché FGA, Pollard JD. Guillain-Barré syndrome, CIDP and other chronic immune-mediated neuropathies. Curr Opin Neurol 1998;11:497-513.

15. Asbury AK and Cornblath DR. Assessment of current diagnostic criteria for Guillain-Barré Syndrome. Ann Neurol 1990;27(suppl):S21-S24.

16. Ad Hoc subcommittee of the American Academy of Neurology AIDS task force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

17. Miescher GC, Steck AJ. Paraproteinaemic neuropathies. Baillieres Clin Neurol 1996;5:219-232.

18. Merkies ISJ, Schmitz PIM, van der Meché FGA, *et al.* Psychometric evaluation of a new sensory scale in immune-mediated polyneuropathies. Neurology 2000;54:943-949.

19. Fünfgeld EW. Vigorimeter: Zur Kraftmessung der Hand und zur Simulationsprüfung. Dtsch Med Wschr 1966;49:2214-2216.

20. Sunderland A, Tinson D, Bradley L, *et al.* Arm function after stroke. An evaluation of grip strength as a measure of recovery and a prognostic indicator. J Neurol Neurosurg Psychiatry 1989;52:1267-1272.

21. Örndahl G, Sellden U, Hallin S, *et al.* Myotonic dystrophy treated with selenium and vitamin E. Acta Med Scand 1986;219:407-414.

22. Rhind VM, Bird HA, Wright V. A comparison of clinical assessments of disease activity in rheumatoid arthritis. Annals Rheum Dis 1980;39:135-137.

23. Toyka KV, Augspach R, Wiethölter H, *et al.* Plasma exchange in chronic inflammatory polyneuropathy: evidence suggestive of a pathogenic humoral factor. Muscle Nerve 1982;5:479-484.

24. Sharrack B, Hughes RAC. Scale development and Guy's neurological disability scale. J Neurol 1999;246:226.

25. Desrosiers J, Bravo G, Hébert R, *et al.* Normative data for grip strength of elderly men and women. Amer J Occup Ther 1995;49:637-644.

26. Fike ML, Rousseau E. Measurement of adult hand strength: a comparison of two instruments. Occup Ther J Res 1982;2:43-49.

27. Robertson A, Deitz J. A description of grip strength in pre-school children. Amer J Occup Ther 1988;42:647-652.

28. Jones E, Hanly JG, Mooney R, *et al.* Strength and function in the normal and rheumatoid hand. J Rheumatol 1991;18:1313-1318.

29. American society of hand therapists. Clinical assessment recommendations, 2nd ed, Chicago, Author, 1992.

30. Dwyer J, Feinleib M. Introduction to statistical models for longitudinal observation. In: Dwyer J, Feinleib M, Lippert P, Hoffmeister H, eds. Statistical models for longitudinal studies of health. New York, Oxford University Press, 1992: 3-48.

31. Nunnally JC. Psychometric theory. New York: McGraw Hill, 1978.

32. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989;27:S178-S189.

33. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. Med Care 1990;28:632-642.

34. Liang MH, Larson MG, Cullen KE, *et al.* Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis Rheum 1985;28:542-547.

35. Cohen J. Statistical power analysis for the behavioural sciences. New York: Lawrence Erlbaum Associates, 1988.

36. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. J Chron Dis 1987;40:171-178.

37. Pham BA, Cranney A, Boers M, *et al.* Validity of area-under-the-curve analysis to summarize effect in rheumatoid arthritis clinical trials. J Rheumatol 1999;26:712-716.

38. Stucki G, Liang MH, Fossel AH, *et al.* Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. J Clin Epidemiol 1995;48:1369-1378.

39. Wright JG, Young NL. A comparison of different indices of responsiveness. J Clin Epidemiol 1997;50:239-246.

40. Atroshi I, Johnsson R, Sprinchorn A. Self-administered outcome instrument in carpal tunnel syndrome. Acta Orthop Scand 1998;69:82-88.

41. Redelmeier DA, Guyatt GH, Goldstein RS. Assessing the minimal important difference is symptoms: a comparison of two techniques. J Clin Epidemiol 1996;49:1215-1219.

Chapter 12

Connecting impairment, disability, and handicap in immune-mediated polyneuropathies

Merkies ISJ[1], Schmitz PIM[2], van Doorn PA[1], Samijn JPA[1], van der Meché FGA[1], for the Inflammatory Neuropathy Cause and Treatment (INCAT) group[3]

From the Department of [1]Neurology, [2]Department of Statistics - Daniel den Hoed Cancer Centre, Erasmus University Medical Centre Rotterdam, dr. Molewaterplein 40, 3015 GD, Rotterdam, The Netherlands. [3]See Chapter 2 for INCAT-members.

Submitted for publication

Abstract

Background: In the postulated framework by the World Health Organization (WHO) – the 'International classification of impairments, disabilities, and handicaps (ICIDH)' – various levels of outcome are suggested to be associated with each other. However, others have criticised the ICIDH, stating that it only represents a general, non-specific relationship between its entities.
Objectives: To examine the significance of the ICIDH in immune-mediated polyneuropathies.
Methods: Four impairment measures (Fatigue severity scale, MRC sumscore, "Incat" sensory sumscore, grip strength with the Vigorimeter), five disability scales (nine-hole peg test, ten meters walking test, an overall disability sumscore [ODSS], Hughes' functional grading scale, Rankin scale), and a handicap scale (Rotterdam 9 items handicap scale [RIHS9]) were assessed in 113 clinically stable patients (83 who experienced Guillain-Barré syndrome (GBS), 22 with chronic inflammatory demyelinating polyneuropathy (CIDP), 8 patients with a gammopathy related polyneuropathy). Regression analyses, with backward and forward stepwise strategies, were performed to determine the correlation between the various levels of outcome (disability on impairment, handicap on impairment, handicap on disability, handicap on impairment plus disability).
Results: Impairment measures explained a substantial part of disability (R^2=0.64) and about half of the variance in handicap (R^2=0.52). Disability measures demonstrated a stronger association with handicap (R^2=0.76). Combining impairment and disability scales accounted for 77% of the variance in handicap (RIHS9) scores.
Conclusion: In contrast with some literature suggestions, support for the ICIDH-model is found in the current study, because significant associations have been demonstrated between the various levels of outcome in patients with immune-mediated polyneuropathies. Future studies are, however, required to examine other possible contributors to patients' deficits in daily life in these conditions.

Introduction

In 1980, The World Health Organization (WHO) described its 'International classification of impairments, disabilities, and handicaps (ICIDH)' model staging the consequences of an underlying pathology (1). In this model, an association between the various dimensions was suggested. The WHO defined the disability level as a "a reflection of the consequences of impairment in terms of functional performance and activity..." (1). Also, the handicap level was described as "a disadvantage for a given individual, resulting from impairment or a disability that limits or prevents the fulfilment of a role that is normal..." (1). Collectively, the different levels of outcome are referred to as "disablement".
Impairment and disability measures might be logical in their use and have been the cardinal targets for physicians to assess outcome in general medicine. However, measuring handicap should be more performed, particularly in patients with chronic conditions or diseases with long-term impact on one's life, because handicap formerly represents the end-stage of the common disablement pathway (1, 2). Despite the conceptual advance of the ICIDH model, others have criticised its concept. In particular, it was argued that the suggested association between impairment, disability, and handicap, represented only a general, non-specific, relationship (3,4). A disappointing association between the ICIDH levels was also demonstrated in cardio-pulmonary conditions (5,6).
The aim of this study was to evaluate the possible linkage between impairment, disability, and handicap in patients with sensory-motor immune-mediated polyneuropathies, using a set of scales that covered the greatest part of all ICIDH levels. The ultimate goal was to determine the proportion of handicap variance explained by the combined impairment and disability measures. It was believed that these evaluations would provide the knowledge of how these conditions might have a long-term influence on one's life. The strength of the relationship between items representing various levels of outcome would also provide the extent to which physicians might use an index as a proxy to measure another level of clinical deficit. As an example, grip strength (an impairment measure) was suggested to be an indirect indicator of arm disability, because it demonstrated a moderate to good association with an arm disability scale in immune-mediated polyneuropathies (7).

Patients and Methods

Patients
One hundred and thirteen patients (83 GBS, 22 CIDP, 8 MGUSP) with a stable clinical condition were recruited from the Rotterdam immune-mediated polyneuropathy databank and the Dutch GBS study group (stable group). Patients with GBS, CIDP or MGUSP were recruited, since it was argued that these disorders represent parts of a continuum regarding their neuromuscular dysfunction pattern (8). The selected patients still had residual symptoms or signs due to their illness, representing a broad range of disability. Nine CIDP patients required interval treatment ranging from weeks to months, with intravenous immunoglobulin (IVIg). With this therapy their clinical condition has been stable for more than 6 months. Six patients with MGUSP (three with IgG, two with IgM, and one patient with IgG+IgM) had an associated demyelinating polyneuropathy with minor concurrent axonal damage in three. An axonal polyneuropathy was diagnosed in the remaining two

patients with MGUSP (one IgA and one IgG type). All GBS and CIDP patients met the international criteria for their illness (9,10). The diagnosis MGUSP was established after excluding all possible causes for the gammopathy and polyneuropathy (11).

Assessment tools/scales

The following scales were selected by a panel of 13 expert neurologists, all members of the Inflammatory Neuropathy Cause And Treatment (INCAT) group – a collaborating force of European neurologists with special interest in neuro-immunological illnesses. The scales were selected taking as much as into account the clinical spectrum of sensory-motor immune-mediated polyneuropathies (8). Moreover, most of these scales have been applied in various immune-mediated polyneuropathy studies.

The recently validated Dutch version of the *Fatigue Severity Scale* (FSS) was used to assess fatigue (12,13). The FSS is a brief and simple self-assessed questionnaire containing nine items with answers ranging from 1 ("strongly disagree") to 7 ("strongly agree") for each inquiry. The mean score of the 9 inquiries ranges from 1 ("no signs of fatigue") to 7 ("most disabling fatigue") (12,13).

The *MRC sumscore* is a summation of the Medical Research Council grades (range: 0 - 5) given in full numbers of the following muscle pairs: upper arm abductors, elbow flexors, wrist extensors, hip flexors, knee extensors, and foot dorsal flexors. The MRC sumscore ranges from 0 ("total paralysis") to 60 ("normal strength") (14). The subdivision MRC-arms (range: 0 -30) and MRC-legs (range: 0 - 30) were also incorporated separately in the univariate regression analyses.

The *"INCAT" sensory sumscore* (ISS) was recently introduced and extensively evaluated in terms of its clinimetric soundness in patients with immune-mediated polyneuropathies (15). In brief, this sensory scale comprises pin-prick and vibration sense plus a two-point discrimination value in the arms and legs and ranges from 0 ("normal sensation") to 20 ("most severe sensory deficit") (15). The sensory modalities representing the ISS were also analysed separately in the univariate regression analyses to determine their impact on disability and handicap (pin-prick arm + leg, range: 0 [no deficit] – 8 [maximum deficit]; vibration arm + leg, range: 0 [no deficit] – 8 [maximum deficit]; 2-point discrimination, range: 0 [no deficit] – 4 [maximum deficit]).

The *Vigorimeter* (VM) (Martin, Tuttlingen, Germany) is a hand-held dynamometer to measure grip strength (7). The medium sized bulb was applied in the selected patients. The pressure in the bulb is registered on a manometer via a rubber junction tube and expressed in kiloPascals (kPa; range: 0 - 160) (7).

The Modified *Hughes' functional grading scale* (*f-score*) assesses the functional ability of patients with a strong emphasis on mobility and ranges from 0 (no symptoms or signs) to 5 (requiring artificial ventilation for at least part of the day) (14).

The *Overall Disability sumscore (ODSS)* is composed by a recently published arm and leg disability scale with a total score ranging from 0 ("no signs of disability") to 12 ("most severe disability score") (16). The ODSS comprises a good functional description of the arms and legs in a checklist form suitable for interviewing patients. Daily arm activities like dressing upper part of the body, doing and undoing buttons and zips, washing and brushing hair, using a knife and fork and turning a key in a lock are scored as being "not affected", "affected but not prevented" or "prevented". Subsequently, these results are translated into an arm grade (score range: 0 [normal arm abilities] to 5 [severe symptoms and signs in both

arms preventing all purposeful movements]). The leg scale highlights problems regarding walking taking into account the use of a device. The results are also translated into a leg grade (score range: 0 [walking is not affected] to 7 [restricted to wheelchair or bed most of the day, preventing all purposeful movements of the legs]) (16). The selected arm and leg disability scales are subsets of a more comprehensive Guy's neurological disability scale (16). The ODSS, its arm disability scale and leg disability scale were separately examined in the univariate regression analyses to determine their association with the handicap level.
The *Rankin scale* has been primarily used in patients with stroke (17). The grades of this scale range from: 0 (no symptoms at all) to 5 (severe disability, bedridden, incontinent, and requiring constant nursing care and attention) (17).
The *Nine-hole peg test* and the *Ten meters walking test* were also applied in all patients to measure 'focal disability' (18-20).
The *Rotterdam 9 Items handicap scale* (*RIHS9*) was recently constructed using the international criteria for development of outcome measures (21,22). The World Health Organisation handicap scale was used as a framework for the construction of the RIHS9, taking into account most of the postulated handicap dimensions by this organisation (1). The RIHS9 comprises 9 inquiries (mobility indoors, mobility outdoors, kitchen tasks, domestic tasks indoors, domestic tasks outdoors, leisure activities indoors, leisure activities outdoors, able to drive a car/go by bus/ride a bicycle, and able to work/study) with raw answers ranging from 1 ("unable to fulfil a task/activity") to 4 ("complete fulfilment of a task/activity"). Since not all items are necessarily applicable to all patients, an answer "0" ("not applicable") was added to all inquiries. To avoid the formation of various subgroups of patients with different amounts of applicable items (for example: a subgroup with 9 applicable items [raw score-range: 9-36], another with 8 [raw score-range: 8-36], etcetera), all initial raw scores were multiplied by 9/(9-number of not applicable items). Hence, the final RIHS9 score-range was independent of the amount of applicable items and extended from 9 ("unable to fulfil any applicable task or activity") to 36 ("able to fulfil all applicable tasks and activities").

Test procedures
All participants gave informed consent before the study. All measures were obtained in a quiet and comfortably warm room at our outpatient clinic. The assessments were performed in a random order. All participants received instructions on how to fill in the FSS form. For the assessment of strength, a standardised joint and limb position as well as the point at which counter-force was administered was defined before the start of the study and taken at examination of each muscle group (page 205). Sensory modalities were examined in triplicate under the earlier prescribed standard conditions with the patients lying in supine position (15). Grip strength with the Vigorimeter was assessed according to the recommendations by the American Society of Hand Therapists (23). Three grip strength measurements with maximum voluntary contractions for each hand were taken in alternating order. Between each trial a pause of 30 seconds was assigned. The results of three trials for each hand were averaged and considered the grip strength score for that particular hand.
All patients received training in assessing the nine-hole peg test before the start of the study to exclude any training effect. The patients were asked, under the prescribed standard conditions in alternating order for both hands, to pick up nine pegs from a tray at table height and place them as quickly as possible into nine holes in a neighbouring horizontal

board. After this procedure, the pegs were removed as fast as possible (18,19). Patients were also requested to walk ten meters in a straight line at their preferential speed, using whatever aid needed (18,20). Three measures were completed for each of these tests and the corresponding time was recorded at each assessment (in seconds). For each test separately, the mean time of completion was calculated by averaging the three obtained measures.

Statistics

Uni- and multivariate linear regression analyses were performed to determine the association between impairment, disability, and handicap outcome measures. The ODSS was chosen as the disability dependent variable in the studies analysing the association with impairment measures (explanatory variables). The ODSS was chosen because of its comprehensiveness in monitoring disability compared with the f-score and Rankin scale (14,17). The latter two scales assess disability with a strong emphasis towards mobility and do not provide information regarding the arms. The RIHS9 was the dependent variable for the analyses of impairment and disability measures leading to handicap. A transformation of the dependent variables (the ODSS: square root; the RIHS9: quadratic transformation) was done to obtain a normal distribution pattern. Univariate regression studies were primarily performed, striving for the best fit between the dependent and independent variable through systematic evaluation of constructed graphs with linear regressions including a restricted cubic spline function on the independent variable (24). Subsequently, multivariate linear regressions were carried out for the various linkages (disability on impairment, handicap on impairment, handicap on disability, handicap on impairment plus disability), using backward eliminating and forward adding stepwise strategies. The strength of association between the dependent variable and explanatory variable was presented as R^2: the fraction of variance explained by the independent variables from a regression model.

In the multivariate regression models, only the right hand grip strength and nine-hole peg values were included and were presented, because these findings turned out to be similar to the regressions that incorporated the left hand values. All analyses were performed using Stata 6.0 for Windows 95 (Stata Statistical Software: Release 6.0. 702 University Drive East, College station, TX: Stata Corporation 1997). A value of $p < 0.05$ was considered statistically significant.

Results

General aspects. The group of patients (54 females and 59 males) had a mean duration of symptoms till onset of the study of 6.9 years. Seven of these patients were bed bound and fourteen required assistance or a device to walk short distances. The remaining 92 patients could walk without any support by someone. The corresponding values for all scales in these patients are listed in Table 1.

Univariate regression studies of impairment leading to disability. The univariate regression studies are presented in Table 2. The MRC sumscore and grip strength values were the strongest explanatory variables of disability, accounting each for 40-45% of the variance in ODSS scores. The MRC score of the legs had a higher impact compared with the score of the arms (Table 2). The strongest explanatory sensory modality of disability was the '2-point discrimination' (Table 2). Fatigue had a non-significant impact on disability.

Table 1

Basic characteristics of patients with immune-mediated polyneuropathies

Stable group of patients (n = 113; GBS 83, CIDP 22, MGUSP 8)	
Mean age at entry (SD), range [years]	54.3 (15.1), 14 – 84
Mean duration of symptoms till onset of the study (SD), range [years]	6.9 (3.1), 0.5 – 28
Sex distribution [%]	
Males	59 (52)
Females	54 (48)
Mean Fatigue severity scale at entry (SD), range	5.6 (1.4), 1 – 7
Mean MRC sumscore at entry (SD), range	53.3 (7.7), 16 – 60
Mean "INCAT" Sensory sumscore at entry (SD), range	4.4 (4.1), 0 – 15
Mean grip strength values with the Vigorimeter at entry (SD), range [kPa]	
Right hands	66.9 (33.6), 0 – 156
Left hands	64.9 (32.9), 0 – 158
Mean Overall Disability sumscore at entry (SD), range	3.5 (2.2), 0 – 11
Mean f-score at entry (SD), range	1.8 (0.9), 1 – 4
Mean Rankin score at entry (SD), range	2.1 (0.9), 0 – 4
Mean nine-hole peg test at entry (SD), range [in seconds] *	
Right hands	31 (20), 15 – 135
Left hands	33 (25), 16 – 192
Mean ten-meters walking test at entry (SD), range [in seconds] **	10.4 (5.8), 5 – 32
Mean Rotterdam nine items handicap scale at entry (SD), range	29.5 (6), 14 – 36

Legend to Table 1
GBS = Guillain-Barré syndrome; CIDP = chronic inflammatory demyelinating polyneuropathy; MGUSP = polyneuropathy associated with a monoclonal gammopathy of undetermined significance. INCAT = inflammatory neuropathy cause and treatment group. *Five patients could not fulfil the nine-hole peg test. **Seven patients were not able to walk.

Univariate regression studies of impairment leading to handicap. Approximately 1/3 of handicap was explained by the MRC sumscore and grip strength, separately. A lower, but still significant association was obtained between the ISS and RIHS9 values ($R^2 = 0.16$). Again, fatigue did not have a significant impact on handicap (Table 2).

Univariate regression studies of disability leading to handicap. The association between disability measures and handicap were higher compared with the univariate regressions that included impairment measures. The Overall disability sumscore was the strongest explanatory variable of handicap, accounting for 65% of the variance in RIHS9 values. The functional grading scale and the Rankin scale were also both highly associated with handicap. The Nine-hole peg test had the weakest association of all disability measures (Table 2).

Multivariate regression studies. In the Figure, the conceptual framework of the World Health Organisation – 'International classification of impairments, disabilities, and handicaps (ICIDH)' – is presented showing the proportion of variances that were obtained from multivariate regression studies between the various levels of outcome in patients with immune-mediated polyneuropathies. As can be seen, approximately 2/3 of disability (assessed by the ODSS) was explained by impairment measures (R^2=0.64). Fatigue (FSS) was the only impairment measure that did not significantly contribute to this model. All impairment measures remained significantly associated with handicap, accounting for a combined 52% of variance in RIHS9 values (Figure). Disability measures explained 76% of handicap disturbances. The Rankin scale and ODSS were the strongest contributors to this model. The Nine-hole peg test and the f-score were eliminated.

Table 2

Univariate regression studies demonstrating the association between impairment, disability, and handicap outcome measures in immune-mediated polyneuropathies (n=113)

Dependent variable →	Overall disability sumscore		Rotterdam 9 Items handicap scale	
Explanatory variable ↴	R^2	p-value	R^2	p-value
Fatigue severity scale	0.03	0.4	0.05	0.1
MRC sumscore	0.45	< 0.0001	0.35	< 0.0001
MRC arms	0.37	< 0.0001	0.24	< 0.0001
MRC legs	0.43	< 0.0001	0.35	< 0.0001
INCAT Sensory sumscore	0.21	< 0.0001	0.16	0.0003
Pin prick arm + leg	0.14	0.0003	0.11	0.002
Vibration arm + leg	0.10	0.002	0.10	0.004
2 point discrimination	0.19	< 0.0001	0.11	0.002
Grip strength				
Right hands	0.40	< 0.0001	0.31	< 0.0001
Left hands	0.43	< 0.0001	0.39	< 0.0001
Functional grading scale	-	-	0.59	< 0.0001
Overall disability sumscore	-	-	0.65	< 0.0001
Arm disability scale	-	-	0.47	< 0.0001
Leg disability scale	-	-	0.52	< 0.0001
Nine-hole peg test				
Right hands	-	-	0.33	< 0.0001
Left hands	-	-	0.40	< 0.0001
Ten meters walking test	-	-	0.53	< 0.0001
Rankin scale	-	-	0.63	< 0.0001

Combining impairment and disability measures accounted for 77% of the variance in handicap scores. In this model, the Rankin scale, the ODSS, the ten-meter walking test, and fatigue (FSS) were the significant explanatory contributors. The values of sensation (ISS), dexterity (Nine-hole peg test), general strength (MRC sumscore), grip strength (Vigorimeter), and functional grading scale (f-score) were all excluded, since their contribution was not significant.

Adding patients' variables (age, sex, and duration of illness) showed only a minor increase in proportion of variances explanation (impairment to disability: R^2 from 0.64 to 0.67; impairment to handicap: R^2 from 0.52 to 0.57; disability to handicap: R^2 from 0.76 to 0.79; impairment + disability leading to handicap: R^2 from 0.77 to 0.80).

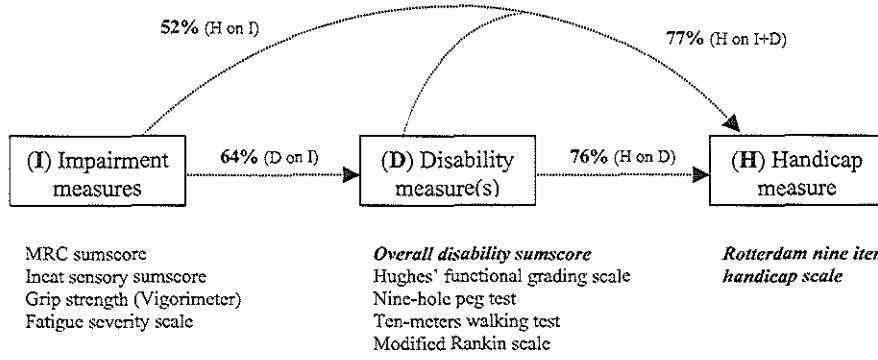
Discussion

In the current study, significant and meaningful associations between the various levels of outcome, as suggested in 1980 by the World Health Organization in their framework – "International classification of impairments, disabilities, and handicaps (ICIDH)", are demonstrated in patients with immune-mediated polyneuropathies (1). These results are in contrast with literature findings that suggested only a marginal, non-specific association between the ICIDH levels of outcome (3-6). Harwood and colleagues also demonstrated the

applicability of the ICIDH model in an outstanding paper addressing the association between these levels in elderly people (2). Moreover, the results in the current study provided the knowledge needed to understand the extent to which for example an impairment measure would contribute to disability or handicap in relation to other impairment variables in these illnesses. Moreover, physicians could use this information to choose a scale that provides outcome at its own level and indirect at another level, depending on the strength of association with other variables. This was also demonstrated in a recent paper addressing grip strength (Vigorimeter) as an adequate monitoring instrument of outcome at the impairment level and indirect at the disability level, because a significant association was observed between grip strength values and an arm disability scale in patients with immune-mediated polyneuropathies (7).

Figure

Multivariate regression studies linking impairment, disability, and handicap in patients with immune-mediated polyneuropathies (n=113)



Legend to Figure
The dependent disability and handicap measures are *italicised.* The proportion of variance for the dependent variables (Overall disability sumscore and Rotterdam nine items handicap scale at the disability and handicap levels, respectively) were calculated and expressed in percentages.

In the current study, only 2/3 of disability was explained by impairment measures we used. The MRC sumscore and grip strength (Vigorimeter) values were the strongest explanatory variables, which is consistent with earlier reports (25,26). Since the association was not absolute, other explanatory variables should be considered as well. In a study evaluating persistent disability in Guillain-Barré syndrome, muscle weakness, sensory dysfunction, fatigue, contractures, and psychological factors were found as persistent items that led to disability (26).

To our knowledge, the current study is the first that analysed the impact of impairment measures leading to handicap in immune-mediated polyneuropathies. However, only 52% of the variance in handicap was explained by impairment variables (Table 2), suggesting the contribution of other explanatory factors. The strongest association was obtained between the combined impairment and disability measures explaining handicap. Almost 80% of

handicap variance was explained by these measures, thus supporting the assumed associations given in the ICIDH model (1). However, the association was not absolute and therefore other explanatory factors should be considered as potential contributors to handicap. Such factors might be pain, psychological items like anxiety, depression, coping mechanism, and motivation, social support, and physical condition in terms of endurance (26-30). The assessment and incorporation of these factors could be cardinal for further improvement in understanding the consequences of immune-mediated polyneuropathies at the various levels of outcome. This is one of the main reasons that the WHO is currently revising its ICIDH model for all diseases and disorders in general (31). Its strive is to construct a more comprehensive model that also integrates personal and environmental factors (31). It should also be stated that the obtained associations in the current study were directly linked to the used scales and might vary if other outcome measures were used.

In conclusion, in the current study, the applicability of the 'International classification of impairments, disabilities, and handicaps' framework is demonstrated in immune-mediated polyneuropathies. Further knowledge is provided in understanding the consequences of these illnesses leading to deficit at the various levels of outcome.

References

1. World Health Organisation. International classification of impairments, disabilities, and handicaps. Geneva: WHO, 1980.

2. Harwood RH, Prince M, Mann A, *et al.* Associations between diagnosis, impairments, disability and handicap in a population of elderly people. Int J Epidemiol 1998;27:261-268.

3. Coulthard-Morris L, Burks JS, Herndon RM. Rehabilitation outcome measures. In: Herndon RM, editor. Handbook of neurologic rating scales. New York: Demos Vermande, 1997, 225-264.

4. Wade DT. Measurements in neurological rehabilitation. Oxford University Press; Oxford, 1992.

5. Cowley AJ, Fullwood LJ, Muller AF, *et al.* Exercise capability in heart failure: is cardiac output important after all? Lancet 1991;337:771-773.

6. Williams SJ, Bury MR. 'Breathtaking': the consequences of chronic respiratory disorder. Int Disabil Stud 1989;11:114-120.

7. Merkies ISJ, Schmitz PIM, Samijn JP, *et al.* Assessing grip strength in healthy individuals and patients with immune-mediated polyneuropathies. Muscle Nerve 2000;23:1393-1401.

8. Hartung H-P, van der Meché FGA, Pollard JD. Guillain-Barré syndrome, CIDP and other chronic immune-mediated neuropathies. Curr Opin Neurol 1998;11:497-513.

9. Asbury AK and Cornblath DR. Assessment of current diagnostic criteria for Guillain-Barré Syndrome. Ann Neurol 1990;27(suppl):S21-S24.

10. Ad Hoc subcommittee of the American Academy of Neurology AIDS task force. Research criteria for diagnosis of chronic inflammatory demyelinating polyneuropathy (CIDP). Neurology 1991;41:617-618.

11. Miescher GC, Steck AJ. Paraproteinaemic neuropathies. Baillieres Clin Neurol 1996;5:219-232

12. Krupp LB, LaRocca NG, Muir-Nash J, *et al.* The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 1989;46:1121-1123.

13. Merkies ISJ, Schmitz PIM, Samijn JPA, *et al.* Fatigue in immune-mediated polyneuropathies. Neurology 1999;53:1648-1654.

14. Kleyweg RP, van der Meché FGA, Schmitz PIM. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré syndrome. Muscle Nerve 1991;14:1103-1109.

15. Merkies ISJ, Schmitz PIM, van der Meché FGA, *et al.* Psychometric evaluation of a new sensory scale in immune-mediated polyneuropathies. Neurology 2000;54:943-947.

16. Sharrack B, Hughes RAC, Soudain S. Guy's neurological disability scale [abstract]. Neurology 1996;243(suppl 2):S32.

17. van Swieten JC, Koudstaal PJ, Visser MC, *et al.* Interobserver agreement for the assessment of handicap in stroke patients. Stroke 1988;19:604-607.

18. Wade DT. Measurements in neurological rehabilitation. Oxford: Oxford University press, 1992.

19. Mathiowetz V, Weber K, Kashman N, *et al.* Adult norms for the nine hole peg test of finger dexterity. Occup Ther J Res 1985;5:24-38.

20. Collen FM, Wade DT, Bradshaw CM. Mobility after stroke: reliability of measures of impairment and disability. Int Disab Studies 1990;12:6-9.

21. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. New York, Oxford university press, 1998.

22. Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chron Dis 1985;38:27-36.

23. American society of hand therapists. Clinical assessment recommendations, Chicago: 2nd ed, 1992.

24. Herndon JE II, Harrell FE jr. The restricted cubic spline hazard model. *Comm Stat Theory Methods* 1990;19:639-663.

25. Molenaar DS, Vermeulen M, de Visser M, *et al.* Impact of neurologic signs and symptoms on functional status in peripheral neuropathies. Neurology 1999;52:151-156.

26. Lennon SM, Koblar S, Hughes RAC, *et al.* Reasons for persistent disability in Guillain-Barré syndrome. Clin Rehabil 1993;7:1-8.

27. Bernsen RAJAM, de Jager AEJ, Schmitz PIM, *et al.* Residual physical outcome and daily living 3 to 6 years after Guillain-Barré syndrome. Neurology 1999;53:409-410.

28. Nicholas R, Playford ED, Thompson AJ. A retrospective analysis of outcome in severe Guillain-Barré syndrome following combined neurological and rehabilitation management. Disabil Rehabil 2000;22:451-455.

29. Pitetti KH, Barrett PJ, Abbas D. Endurance exercise training in Guillain-Barré syndrome. Arch Phys Med Rehabil 1993;74:761-765.

30. de Jager AEJ, Minderhoud JM. Residual signs in severe Guillain-Barré syndrome: analysis of 57 patients. J Neurol Sci 1991;104:151-156.

31. ICIDH-2:International classification of impairments, activities, and participation. A manual of dimensions of disablement and functioning. Beta-1 draft for field trials. World Health Organisation, Geneva, 1997.

Chapter 13

Summary - General discussion - Future perspectives

13.1 The new International Classification of Impairments, Activities, and Participation (ICIDH-2)

13.2 Summary and future perspectives

13.1 The new International Classification of Impairments, Activities, and Participation (ICIDH-2)

Historically, the exponential growth in health care during the 20th century has brought with it the need to focus on the "consequences" of disease/health conditions rather than the disease alone. With the improvement in general medicine, the growing importance of chronic conditions, and the ageing of the population, the "consequences of illnesses" gained in significance because of the life-long management needs. In particular, the functional management of an illness became the goal and the use of outcome measures became the standard for measuring the performance of health care delivery and its effectiveness. The World Health Organisation (WHO) postulated a framework, the international classification of impairments, disabilities, and handicaps (ICIDH), to structure outcome measures used in the evaluation of the consequences of an illness (1). In the current study, various scales and gadgets were chosen and evaluated to represent purely one of these outcome levels. Also, a generic quality of life measure was selected (1,2).
Despite the general acceptance of the ICIDH as a detailed and unified system, some remarks should be addressed. Considerable experience has been gained within the last two decades using this model in all specialties worldwide (1). However, there is a growing necessity to revise the ICIDH in the light of new ideas that has emerged on the consequences of a disease. The graphical representation of the ICIDH model (Chapter 1; figure 2) is helpful in distinguishing between impairment, disability, and handicap concepts, but:

- it does not provide adequate information on the relationship between these concepts;
- it does not incorporate health conditions other than illnesses;
- the arrows linking disease (pathology), impairment, disability, and handicap have occasionally been interpreted as representing a "causal model" and an indication of change over time;
- this representation does not allow movement from handicap and disability back to impairment, and has thus been taken to imply a *unidirectional flow* from impairment to the other concepts;
- it does not adequately reflect the role of the social and physical environment in the consequences of a disease;
- although the original text states that the situation is more complex than a simple linear progression, this statement needs to be made more clearly – the arrows in the graphic presentation must be understood as meaning no more than "may lead to".

Prompted by these observations, the WHO is currently evaluating a new and more comprehensive version of the ICIDH. The draft of this new multidimensional framework, named "the International Classifications of Impairments, Activities, and Participation (ICIDH-2)" integrates the various aspects related to the consequences of a health condition (e.g. a disease, disorder or injury). The current understanding of interactions within the ICIDH-2 is presented in the figure. As can be noted, dimensions have been nominated by "neutral" terms. "Disability" has been replaced by "Activity", based on the activities/limitations of a person in daily life. "Handicap" has been reformulated as "Participation", hence introducing a positive concept for each dimension. This frame illustrates also the dynamic interaction among the various dimensions. The draft of this

ICIDH-2 is available on the Internet (3). In the following, a definition of the various levels is provided.

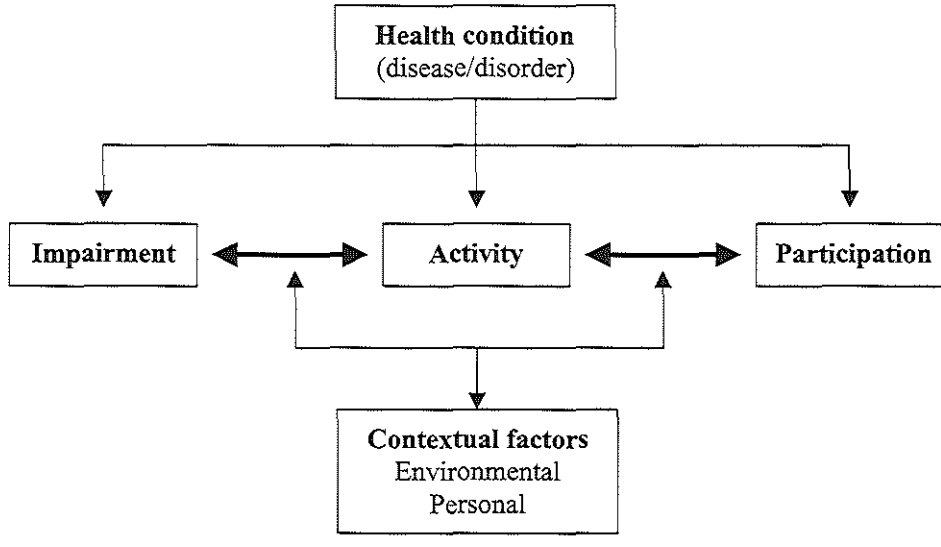
- A *health condition* is an alteration of the health status of an individual that may lead to distress, interference with daily activities, or contact with health services. It may be a disease (acute or chronic), disorder, injury or trauma, or reflect other health related states such as pregnancy, ageing, stress, congenital anomaly, or genetic predisposition

- The *Impairment* dimension is defined as a loss or abnormality of body structure or of a physiological or psychological function. It relates to either body functions or body structures.

- The *Activity* dimension is defined as the nature and extent of functioning at the level of the person. It integrates activities of a person as performed in daily tasks.

- The *Participation* dimension is defined as the nature and extent of a person's involvement in life situations in relation to Impairments, Activities, Health conditions, and Contextual factors. It represents the consequences of health conditions at societal level.

Figure



- *Contextual factors* are defined as the complete background to a person's life and living. These factors include environmental factors and personal factors. Environmental factors are extrinsic to (outside of) the individual (e.g. the attitude of society, architectural characteristics, the legal system). Personal factors are intrinsic to (inside of) the individual and describe on how the consequences of a health status

is experienced (e.g. gender, age, fitness, lifestyle, habits, coping styles, social background, education, life experience).

On the Internet section, various case examples are presented illustrating the interaction and relationship between all these terms (3). Compared with the 1980 ICIDH model, the ICIDH-2 model strives for a more comprehensive description of all aspects possible related to an individual's reaction on handling the consequences of a health condition. More emphasis is directed toward a patient's daily activities and social interactions. The ICIDH-2 is currently being evaluated in various fields and trials and the final draft is expected in 2001.

13.2 Summary and future perspectives

In **chapter 1**, the clinical aspects of the most commonly forms of immune-mediated polyneuropathies (Guillain-Barré syndrome [GBS], chronic inflammatory demyelinating polyneuropathy [CIDP], monoclonal gammopathy of undetermined significance polyneuropathy [MGUSP], multifocal motor neuropathy [MMN]) are presented. It was stated that these disorders form a spectrum with no clear-cut boundaries, despite differences in time course and response to various treatments. A systematic review of all clinical studies in these conditions, published in the English language between January 1988 and January 1999 that included at least ten patients, demonstrated an enormous array of available and applied outcome measures. In particular, many scales have been devised and used to examine general strength, sensory deficit, or functional ability. However, most of these scales have not been evaluated thoroughly in terms of their fulfilment of all clinimetric requirements, such as being easily applicable, valid, reliable, and responsive to changes over time. A comprehensive review of the literature findings is presented in this chapter. Also, relatively little attention has been focused on the consequences of these illnesses, particularly at the disability and handicap levels of outcome, and from the patient's own perspective as captured by quality of life surveys (1,2). It was concluded that no uniformity existed regarding which set of scales represented best the different levels of outcome.
Based on these shortcomings, the Inflammatory Neuropathy Cause And Treatment (*INCAT*) group – a collaborating force of mainly European neurologists with special interest in neuro-immunological disorders – decided to conduct a study in which a set of scales and gadgets was selected and evaluated in terms of their clinimetric properties in immune-mediated polyneuropathies. Two scales were constructed and added to the set of outcome measures and helped in covering the whole clinical range in these conditions. Moreover, the selected scales and gadgets represented various entities at different levels of outcome (impairment, disability, handicap, quality of life), thus providing information on the impact of these disorders from different angles. Their mutual associations were also examined to determine the interactions between the various outcome levels. Also, clinically useful normal values were determined for the selected Martin Vigorimeter and the Rydel-Seiffer graduated tuning fork.

In **chapter 2** clinically useful vibration threshold normative values for the arms and legs are presented for the graduated Rydel-Seiffer tuning fork. Also, its simplicity, and validity were demonstrated by correlation with the Vibrameter. A total of 198 healthy controls, stratified for sex and age, were examined at the index finger, styloid proces, hallux, and internal malleolus. Acceptable sensitivity and specificity were demonstrated for the tuning fork when evaluating a group of mildly affected patients with various forms of polyneuropathy. The tuning fork detected more abnormalities in the legs compared with the results obtained with the Vibrameter.
The following should be addressed regarding the sites of examination and obtained vibration threshold normative values for the Rydel-Seiffer tuning fork. A randomly selected group of 40 (18 men; 22 women; mean age 58.4 [SD 18.6] years) healthy controls was also clinically examined to obtain vibration values using the Rydel-Seiffer tuning fork at the medial humerus epicondyle, acromioclavicular joint, patella, and anterior superior iliac spine. The aim was to determine whether the obtained values would be different than the presented

clinically useful normative values for the more distal sites of examination with this fork. The obtained proximal values in this subgroup of healthy controls demonstrated only slight differences (slightly higher values; data not published) with the presented normative values for the distal sites of examination. Hence, to strive for clarity, the published vibration threshold normative values were extrapolated to the more proximal regions of the extremities.
Although not presented in chapter 2, the influence of length on the vibration values in the regression analyses was also examined. Length demonstrated an inverted correlation with the recruited vibration threshold values at all sites of examination, but particularly in the legs. However, its influence on the final vibration value was negligible compared with age. The latter remained the strongest predictor of vibration sense values.

In **chapter 3**, the reliability and responsiveness of the Rydel-Seiffer tuning fork were demonstrated in patients with immune-mediated polyneuropathies. Thus, all clinimetric requirements were demonstrated for this fork, by combining these results with literature findings. The incorporation of this easily applicable instrument in routine neurological examination was therefore suggested.

Chapter 2 and 3 partially form the basis for the construction and use of the *INCAT* sensory sumscore (ISS) in patients with immune-mediated polyneuropathies, as presented in **chapter 4**. The use of the ISS was suggested to strive for uniformity in assessing sensory deficit in these disorders. This scale was easily applicable and required only a median time of 4.5 minutes for completion. Acceptable validity, reliability, and responsiveness were also demonstrated. Despite its fulfilment of all clinimetric requirements, further improvement of this scale is possible. In particular, normal values on the two-point discrimination test could be determined, thus refining the final outcome of this scale. In a paper from Japan, an increase in distance of two-point discrimination was seen with advancing age. Women had slightly higher mean values compared with men (4). Also, future studies should determine whether the ISS captures changes over time in patients with MGUSP, since the course in these patients is more indolent compared with patients with GBS and CIDP.

In **chapter 5**, clinical useful normative grip strength values for the hand-held Vigorimeter were presented depending primarily on age and sex. A total of 530 healthy individuals were investigated, with and age ranging from 5 to 93 years and stratified for age and sex. Also, all clinimetric requirements were demonstrated for this easily applicable instrument. In the longitudinally examined patients, higher and more normal grip strength values were obtained over time. Moreover, the Vigorimeter correlated significantly with an arm disability scale and provided, therefore, outcome at the impairment level and indirect at the disability level.

In **chapter 6**, a disability scale was introduced and evaluated in patients with immune-mediated polyneuropathies. This overall disability sumscore comprises a good functional description of the arms and legs in a checklist form suitable for interviewing patients. The arm component addresses daily arm activities and the leg subscale highlights problems regarding walking taking into account the use of a walking device or support. The comprehensiveness of this scale is in contrast with the widely used Hughes' functional

grading scale (f-score) and Rankin scale, which have a strong emphasis towards mobility and do not provide information regarding the arms. The validity, reliability, and responsiveness of the overall disability sumscore were demonstrated. Also, impairment disturbances like general weakness and sensory deficit leading to disability were better monitored by the overall disability sumscore compared with the f-score and Rankin scale.

In **chapter 7**, a prospective open study was performed evaluating outcome at the impairment and disability levels in patients with chronic motor neuropathies not improving after conventional therapies who were treated with interferon-β1a. Slight improvement was noted in all patients on this drug, demonstrating more general strength, and better dexterity and ambulation. Subjectively, performing daily activities improved also. However, the modified Rankin scale did not capture improvement in all patients. This was most probably due to the low responsiveness of this scale to capture relevant changes over time. Others using the Rankin scale also reported similar findings (5).

In **chapter 8**, the development and clinimetric evaluation of a new handicap scale, the Rotterdam 9 items handicap scale, was described. Its construction was accomplished based on the international guidelines for scales construction using the World Health Organisation's handicap dimensions as a framework. Extensive literature review for potential items, suggestions and judgements of the selected items by patients who suffered from immune-mediated polyneuropathies, and expert opinions by a group of neurologists, were the main phases in its construction. Subsequently, all clinimetric requirements were demonstrated in patients with immune-mediated polyneuropathies. A median time of 3.5 minutes was needed to complete this handicap scale. The use is this scale was therefore suggested for monitoring outcome at the handicap level in these conditions. Future studies are however required to determine which handicap measure will ultimately be more feasible in these conditions.

In **chapter 9**, the impact of fatigue on functionality and quality of life is presented in a group of clinically stable patients with immune-mediated polyneuropathies. It was stated that fatigue might have been under-recognised by neurologists and other disciplines, because attention is generally directed towards weakness and sensory deficit in these conditions. This is the first paper reporting on fatigue in these conditions. Fatigue turned out to be a prominent and highly disabling symptom in these patients. At least 80% of the patients at study were severely fatigued and this was irrespective of other clinical parameters such as weakness, sensory deficit, and duration of illness. Remarkably, fatigue scores in patients with GBS were significantly associated with the SF-36 socio-emotional dimensions but not with the physical subscales. Like in the paper by Bernsen *et al.*, it was concluded that the psychosocial functioning of patients with GBS was seriously affected, even when the patients reached a complete physical (weakness and sensory deficit) recovery or showed only mild residual signs (6). Conversely, fatigue values in patients with CIDP of MGUSP were significantly related to the physical but not the socio-emotional SF-36 dimensions. It was assumed that these patients were more preoccupied with the potential threat of changes in their physical status, making them prone to relate fatigue to their physical condition.
Despite these results, there all still many questions unanswered. What is the pathophysiological mechanism of fatigue in these conditions? Are there therapeutic options for fatigue? Future studies of patients with these illnesses should focus on these items.

Krupp and Pollina described mechanisms and management strategies of fatigue in progressive neurologic disorders (7). In this review paper, various possible pathophysiological mechanisms, contributing to fatigue were briefly highlighted with particular interest for the reported treatment options thus far (7). Because great similarity in patho-immunological mechanism has been postulated between immune-mediated polyneuropathies and multiple sclerosis (MS), it is suggested that attention should be focused on the treatment strategies of fatigue shown to be effective in patients with multiple sclerosis, such as exercise programs to combat deconditioning and pharmacological therapy with amantadine (8-11). These therapeutic options could be extrapolated for evaluation in patients with immune-mediated polyneuropathies suffering from severe fatigue. Other possible pathophysiologic associated factors with fatigue (for example, social network, depression, and sleep disturbances) deserve a systematic investigation as well (12-14).
Despite being user-friendly and non-time consuming, valid and reliable, the Fatigue severity scale (FSS) should be used as a screening instrument alone, because it addresses only the dimension of subjective feeling of fatigue. It does not provide information regarding the various ways in which fatigue affects patients' lives. Therefore, a more comprehensive assessment of fatigue should be entailed in future studies using multi-dimensional fatigue scales alongside the FSS. However, despite these limitations, the FSS provided in the current study indirect information by correlation with other measures, such as the SF-36 health survey. Finally, the responsiveness to changes over time of the FSS in longitudinally examined patients should be examined as well.

In **chapter 10**, the evaluation of quality of life using the medical outcome study 36-items health status (SF-36) is described in patients with immune-mediated polyneuropathies (15). SF-36 registered adequately the physical and psychosocial shortcomings in these conditions when compared with the reported normal values for the healthy Dutch community (16,17). Also, acceptable clinimetric requirements were demonstrated for this scale, thereby facilitating its general applicability in these conditions. SF-36 enriched physicians' awareness by providing a more holistic view of the impact of having an immune-mediated polyneuropathy.
In the longitudinally examined patients, gradual improvement was noted at all SF-36 levels during follow-up. It was suggested that an extensive guidance most probably contributed to a better outcome in the longitudinally examined patients. Support for this hypothesis was addressed in a study where a supportive group of individuals surrounding and guiding a patient played an important role in the promotion of general health (18). However, future studies should determine whether various pre-defined forms of supportive social network would lead to different outcomes.
It is of great importance also to evaluate the possible differences in quality of life between patients with GBS, CIDP, and MGUSP. Finally, a feasibility study of various generic quality of life measures in patients with immune-mediated polyneuropathies is also needed to determine which health status fits these conditions best.

Of all clinimetric requirements, responsiveness has been studied the least in general medicine. **Chapter 11** describes a study in which, besides the evaluation of validity and reliability, various known parametrical responsiveness methods and 2 newly created non-parametrical responsiveness techniques were compared using scales that purely represented

either the impairment or disability level of outcome. In general, responsiveness techniques ranked the evaluated scales consistently and enabled the clinician to choose among equally valid and reliable outcome measures. The overall disability sumscore, which provides information on arm and leg functionality, the MRC sumscore, and grip strength as assessed with the hand-held Vigorimeter, were ranked among the best and demonstrated good mutual correlations. The use of these outcome measures is, therefore, primarily recommended in the follow-up of patients with sensory-motor immune mediated polyneuropathies. Also, it was argued that the non-parametrical methods (Schmitz's distribution-free responsiveness score and the Wilcoxon matched pairs signed rank test) increased the correctness when evaluating responsiveness, because many scales demonstrated a non-Gaussian distribution. Moreover, in contrast with literature assumptions, the graphical reflection of responsiveness demonstrated that this entity has a dynamic modulating pattern.
Future studies are required to determine whether a larger group of longitudinally examined patients would give the same rank order of selected scales. Also, discriminative responsiveness between groups of patients (for example, between a treated and a control group) should be investigated (19). Attention should not be focused only on the statistical discriminatory ability of an outcome measure, but also on the concept of its "minimal clinical important difference" (MCID) (20-23). Statistically significant difference indicates whether the hypothesis of no difference can be rejected and the MCID is defined as the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate a change in the patient's functionality (21,22). To date, no effort has been put into defining the MCID for the measures most widely used in studies that examined patients with immune-mediated polyneuropathies. A plausible explanation for this is that, through repeated use, clinicians develop an intuitive sense of the MCID of a measure. For example, neurologists familiar with the MRC sumscore would have no difficulty to specify the clinical significance of a 10% reduction of strength on the MRC sumscore in a newly diagnosed patient with GBS. These clinicians are able to come to certain conclusions having observed a large number of patients, and seeing the changes in function and in clinical course that correspond to the variations in outcome results (21). Hence, despite the lack of objectivity, an extensive clinical experience with a scale or measurement instrument should be considered as a feasible and valid method of determining the significance of changes in applied outcome measure (21). Perhaps, the extensive experience of clinicians could serve as guidance to determine the MCID of widely used measures in immune-mediated polyneuropathies. However, no consensus exists regarding who is to say (the clinician or the patient) what should be considered as minimal clinical importance (20-23). Therefore, as has been demonstrated in rheumatic illnesses for more than a decade, efforts should be directed towards organising consensus meetings on outcome measures in neurology, particularly in the field of immune-mediated polyneuropathies (24). Expert neurologists and researchers on this field should contribute to these meetings, thereby striving for clarification of concepts like minimal clinical important difference.

In **chapter 12**, the linkage between selected impairment, disability, and handicap outcome measures is demonstrated in patients with immune-mediated polyneuropathies. Hence, support for the concept of the 'international classification of impairments, disabilities, and handicaps' is found, since acceptable associations were demonstrated between the various outcome levels (1). Impairment measures explained two-third of disability and about half of

handicap. The MRC sumscore and grip strength values by the Vigorimeter were the strongest explanatory variables in these models. Disability measures explained 76% of handicap with the overall disability sumscore as the strongest explanatory variable. Combining impairment and disability scales accounted for 77% of the variance in handicap. However, since these associations were not absolute, other possible explanatory factors were suggested for future studies like pain, psychological items such as anxiety, depression, coping mechanis, and motivation, social support, and physical condition in terms of endurance. (25-29).

In conclusion, this thesis describes a set of studies in which various selected outcome measures, representing different levels of outcome and covering the whole clinical range in patients with immune-mediated polyneuropathies, were evaluated in terms of their clinimetric requirements. Clinical useful normative values were also provided for the Rydel-Seiffer tuning fork and the hand-held Martin Vigorimeter. Moreover, not only the applicability, validity, and reliability of a measure received attention, but also the concept of being responsive to changes over time was also extensively highlighted. New responsiveness techniques were introduced and the method of calculating the Area-under-the-curve for responsiveness was introduced in neuromuscular disorders. Various symptoms were investigated in these conditions, including fatigue, and the consequences of these disorders on daily activities, at the societal functional level, and quality of life were examined. Hence, this thesis could be considered as a basis for further analyses in the clinimetric field in patients with immune-mediated polyneuropathies.

References

1. World Health Organization. International classification of impairments, disabilities, and handicaps. Geneva: WHO, 1980.

2. Aaronson NK. Quality of life: What is it? How should it be measured? Oncology 1988;2:69-76.

3. World Health Organization. International classification of impairments, activities, and participations. A manual of dimensions of disablement and functioning. Beta-2 draft for field trials. Geneva: 1999. (www.who.ch/icidh)

4. Shimokata H, Kuzuya F. Two-point discrimination test of the skin as an index of sensory ageing. Gerontology 1995;41:267-272.

5. Thompson N, Choudhary P, Hughes RAC, *et al.* A novel trial design to study the effect of intravenous immunoglobulin in chronic inflammatory demyelinating polyradiculoneuropathy. J Neurol 1996;243:280-285.

6. Bernsen RAJAM, Jacobs HM, de Jager AEJ, *et al.* Residual health status after Guillain-Barré syndrome. J Neurol Neurosurg Psychiatry 1997;62:637-640.

7. Krupp LB, Pollina DA. Mechanisms and management of fatigue in progressive neurological disorders. Curr Opin Neurol 1996;9:456-460.

8. Giovannoni G, Hartung HP. The immunopathogenesis of multiple sclerosis and Guillain-Barré syndrome. Curr Opin Neurol 1996;9:165-177.

9. Murray TJ. Amantadine therapy for fatigue in multiple sclerosis. Can J Neurol Sci 1985;12:251-254.

10. Rosenberg GA, Appenzeller O. Amantadine, fatigue, and multiple sclerosis. Arch Neurol 1988;45:1104-1106.

11. Petajan JH, Gappmaier E, White AT, *et al.* Impact of aerobic training on fitness and quality of life in multiple sclerosis. Ann Neurol 1996;39:432-441.

12. Krupp LB, LaRocca NG, Muir J, *et al.* A study of fatigue in systemic lupus erythematosus. J Rheumatol 1990;17:1450-1452.

13. McKinley PS, Ouellette SC, Winkel GH. The contributions of disease activity, sleep patterns, and depression to fatigue in systemic lupus erythematosus. A proposed model. Arthritis Rheumatism 1995;38:826-834.

14. Lichstein KL, Means MK, Noe SL, *et al.* Fatigue and sleep disorders. Behav Res Ther 1997;35:733-740.

15. Ware Jr. JE, Snow KK, Kosinski M, *et al.* SF-36 health survey. Manual and interpretation guide. Boston: The Health Institute, New England Medical Center, 1997.

16. Aaronson NK, Muller M, Cohen PDA, *et al.* Translation, validation, and norming of the Dutch language version of the SF-36 health survey in community and chronic disease populations. J Clin Epidemiol 1998;51:1055-1068.

17. Ware Jr. JE, Gandek B, Kosinski M, *et al.* The equivalence of SF-36 summary health scores estimated using standard and country-specific algorithms in 10 countries: results form the IQOLA project. J Clin Epidemiol 1998;51:1167-1170.

18. Krol B, Sanderman R, Suurmeijer TPBM. Social support, rheumatoid arthritis and quality of life: concepts, measurements and research. Patient Education and Counseling 1993;20:101-120.

19. Guyatt G, Walter S, Normal G. Measuring change over time: Assessing the usefulness of evaluative instruments. *J Chron Dis* 1987;40:171-178.

20. Redelmeier DA, Guyatt GH, Goldstein RS. Assessing the minimal important difference in symptoms: a comparison of two techniques. J Clin Epidemiol 1996;49:1215-1219.

21. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Controlled Clin Trials 1989;10:407-415.

22. Wells GA, Tugwell P, Kraag GR, *et al.* Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. J Rheumatol 1993;20:557-560.

23. Wright JG. The minimal important difference: Who's to say what is important? J Clin Epidemiol 1996;49:1221-1222.

24. Annonymous. Omeract conference on outcome measures in rheumatoid arthritis clinical trials. J Rheumatol 1993;20:525-590.

25. Lennon SM, Koblar S, Hughes RAC, *et al.* Reasons for persistent disability in Guillain-Barré syndrome. Clin Rehabil 1993;7:1-8.

26. Bernsen RAJAM, de Jager AEJ, Schmitz PIM, *et al.* Residual physical outcome and daily living 3 to 6 years after Guillain-Barré syndrome. Neurology 1999;53:409-410.

27. Nicholas R, Playford ED, Thompson AJ. A retrospective analysis of outcome in severe Guillain-Barré syndrome following combined neurological and rehabilitation management. Disabil Rehabil 2000;22:451-455.

28. Pitetti KH, Barrett PJ, Abbas D. Endurance exercise training in Guillain-Barré syndrome. Arch Phys Med Rehabil 1993;74:761-765.

29. de Jager AEJ, Minderhoud JM. Residual signs in severe Guillain-Barré syndrome: analysis of 57 patients. J Neurol Sci 1991;104:151-156.

Samenvatting

In dit proefschrift worden diverse geselecteerde uitkomstmaten, bestaande uit beoordelingsschalen en onderzoeksinstrumenten, onderzocht op hun klinische toepasbaarheid bij patiënten met een immuun-gemedieerde polyneuropathie. Immuun-gemedieerde polyneuropathieën vormen een spectrum van neurologische ziekten die als overeenkomst een aantasting van zenuwen in de armen en benen hebben door een stoornis in het afweersysteem. In **hoofdstuk 1** wordt de basis van dit onderzoek stapsgewijs aangedragen. Als eerste worden de klinische verschijnselen beschreven van de meest voorkomende immuun-gemedieerde polyneuropathieën, te weten het Guillain-Barré syndroom (GBS), de chronische vorm hiervan – chronische inflammatoire demyeliniserende polyneuropathie (CIDP), de polyneuropathie ten gevolge van een afwijkend eiwit in het bloed; de zogenaamde monoclonale gammopathie van onduidelijke betekenis gerelateerde polyneuropathie (MGUSP), en de multifocale motore neuropathie (MMN). Deze ziektebeelden worden gekenmerkt door zwakte en/of gevoelsstoornissen. Een systematische evaluatie van de literatuur werd uitgevoerd waarbij alle in de Engelse taal gepubliceerde klinische studies bij deze aandoeningen tussen januari 1988 tot januari 1999 werden onderzocht. Vervolgens werd gekeken of de gebruikte uitkomstmaten voldeden aan de algemene vereisten zoals *validiteit* (meten wat het hoort te meten), *betrouwbaarheid* (dezelfde uitkomst bij herhaaldelijke metingen) en of veranderingen bij de patiënten in de tijd, bijvoorbeeld ten gevolge van een bepaalde behandeling, geregistreerd werden door de gebruikte uitkomstmaten (*responsief* zijn). Deze vereisten worden overkoepeld door de wetenschappelijke term "**klinimetrie**": de leer die vanuit het klinisch handelen uitkomstmaten onderzoekt op hun toepasbaarheid. De meest gebruikte schalen en instrumenten bleken overwegend gericht te zijn op het meten van de kracht, het gevoel, of het dagelijkse functioneren van patiënten. Verder bleken de meeste uitkomstmaten niet of incompleet te voldoen aan de algemene regels van validiteit en betrouwbaarheid. Ook was er opvallend weinig onderzocht voor wat betreft de responsiviteit van schalen bij deze aandoeningen. Sommige schalen waren verder dusdanig opgesteld dat de uiteindelijke uitkomst niet goed te begrijpen was, doordat bijvoorbeeld verschillende meet-kwaliteiten bij elkaar gevoegd werden die niet samen geëvalueerd konden worden (bijvoorbeeld het meten en samenvoegen van gevoel + lopen + reflexen). Derhalve voldeden deze schalen niet aan de internationale richtlijnen zoals opgesteld door de wereld-gezondheidszorg-organisatie. Deze organisatie heeft een raamwerk gemaakt waarbij metingen bij patiënten gedaan konden worden op verschillende niveau's, te weten: impairment (stoornissen aan het lichaam, bijvoorbeeld zwakte), disability (stoornissen in de zelfzorg, bijvoorbeeld moeite met knopen dicht-/openmaken) en handicap (stoornissen in het functioneren op sociaal vlak, bijvoorbeeld niet kunnen werken).
Op basis van deze bevindingen werd als tweede stap een groep van uitkomstmaten gekozen om het gehele spectrum van patiënten met immuun-gemedieerde polyneuropathieën te evalueren. Naast impairment, disability en handicap schalen werd ook een algemene kwaliteit van leven schaal gekozen. Twee schalen werden volgens internationale richtlijnen geconstrueerd en toegevoegd aan deze set. Voorts werden normaalwaarden verzameld voor de geselecteerde Martin Vigorimeter (een knijpkracht instrument) en de Rydel-Seiffer stemvork, welke voorzien is van een schaalverdeling voor kwantitatieve meting van de vibratiezin.

Deze studies werden uitgevoerd met de medewerking van en namens de "Inflammatory Neuropathy Cause And Treatment - *INCAT*" groep, een samenwerkingsverband van prominente Europese neurologen met speciale belangstelling voor immuun-gemedieerde neurologische aandoeningen.

In **hoofdstuk 2** worden normaalwaarden voor de vibratiezin van de armen en benen beschreven die onderzocht zijn met de Rydel-Seiffer gecalibreerde stemvork. Hierbij werden in totaal 198 gezonde mannen en vrouwen, gelijk verdeeld voor leeftijd en geslacht, onderzocht. Dit instrument werd ook gebruikt bij een groep patiënten met een lichte vorm van polyneuropathie om het gebruik in de praktijk te toetsen. Deze stemvork bleek erg simpel in het toepassing te zijn en vertoonde verder een goede overeenkomst met vibratiedrempel-waarden zoals verkregen met een electronisch vibratie meet-apparaat. Hierdoor werd ook de validiteit van de Rydel-Seiffer stemvork bewezen. Zoals eerder aangegeven, moet een schaal of instrument valide zijn, maar ook betrouwbaar en responsief op klinische veranderingen van patiënten in de tijd. Een goede betrouwbaarheid en responsiviteit van de Rydel-Seiffer stemvork worden beschreven in **hoofdstuk 3** bij patiënten met immuun-gemedieerde polyneuropathieën. Hoofdstukken 2 en 3 zijn complementair aan elkaar en beschrijven gezamenlijk de gehele klinimetrische evaluatie van deze stemvork.

De klinimetrische evaluatie van de Rydel-Seiffer stemvork vormde deels de basis voor de constructie van een algemene schaal om gevoelskwaliteiten te meten. Deze wordt in **hoofdstuk 4** beschreven. In deze schaal worden diverse kwaliteiten van het gevoel gemeten, te weten pijnzin, vibratiezin en 2-punts discriminatie. Deze schaal werd klinimetrisch geëvalueerd waarbij alle vereisten vervuld werden. Circa 4.5 minuut was er nodig om deze uitkomstmaat te meten. De constructie van deze schaal heeft als uiteindelijke doel het streven naar uniformiteit bij het meten van gevoelsstoornissen bij patiënten met (immuun-gemedieerde) polyneuropathieën.

In **hoofdstuk 5** worden normaalwaarden voor de knijpkracht gepresenteerd die met de Martin Vigorimeter zijn verkregen door 530 gezonde mannen en vrouwen te onderzoeken variërend in leeftijd van 5 tot 93 jaar. Deze gezonde mensen waren gelijk verdeeld voor wat betreft geslacht en leeftijd. Alle klinimetrische vereisten (validiteit, betrouwbaarheid, responsiviteit) werden verkregen in patiënten met immuun-gemedieerde polyneuropathieën en dit handige instrument was ook makkelijk toepasbaar. De nieuwe GBS/CIDP patiënten die in de tijd werden onderzocht vertoonden meer hoge en vaker normale knijpkracht waarden in vergelijking met eerdere bevindingen. De Vigorimeter bleek ook een goede correlatie te vertonen met een arm-zelfzorg schaal en voorzag derhalve ook informatie over het functioneren van de armen bij deze patiënten.

In **hoofdstuk 6** wordt een algemene zelfzorg-schaal geintroduceerd en geëvalueerd in patiënten met immuungemedieerde polyneuropathieën. Deze schaal bestaat uit een goede beschrijving van de arm en been functies in een handige vragenlijst-vorm welke geschikt is voor het interviewen van patiënten. Dagelijkse activiteiten van de armen zoals het aankleden, mes en vork gebruiken, haren wassen en kammen alsmede problemen verbonden aan het lopen worden goed beschreven. De klinimetrische vereisten van deze zelfzorg schaal

worden ook in dit hoofdstuk beschreven. Ook bleken stoornissen aan het lichaam die tot zelfzorg stoornissen kunnen lijden meer adequaat door deze schaal geregistreerd te worden in tegenstelling tot andere zelfzorg schalen.

In **hoofdstuk 7** wordt een klinische studie beschreven waarbij vier patiënten met een puur motore (alleen zwakte) neuropathie behandeld werden met een bepaald medicijn (Rebif®: interferon-beta) dat het afweersysteem beinvloedt. Deze patiënten bleken niet te reageren op gangbare behandelingsmethoden. Het effect van dit middel werd gemeten op het niveau van impairment (stoornissen aan het lichaam) en disability (zelfzorgstoornissen). Enige verbetering in kracht en functioneren in het dagelijks leven werden bij all patiënten waargenomen. Eén patiënt vertoonde een duidelijke verbetering. Gesuggereerd werd dat dit middel mogelijk een gunstig effect heeft op de klinische toestand van patiënten met een motore neuropathie die niet blijken te reageren op gangbare behandelingsmethoden.

De ontwikkeling en klinimetrische evaluatie van een nieuwe 'handicap' schaal, de Rotterdam 9 items handicap schaal, worden beschreven in **hoofstuk 8** in patiënten met immuun-gemedieerde polyneuropathieën. De opzet van deze schaal was gebaseerd op de internationale criteria voor uitkomstmaten ontwikkeling waarbij de handicap dimensies, beschreven door de wereld-gezondheidszorg-organisatie, als leidraad gebruikt werden. Uitvoerige literatuur studie op zoek naar potentiële items, suggesties en beoordeling van de geselecteerde items door patiënten met immuun-gemedieerde polyneuropathieën en neurologen met speciale belangstelling voor deze aandoeningen, vormden de belangrijkste steunpiralen voor de constructie van de uiteindelijke handicap schaal. Vervolgens werden alle klinimetrische vereisten aangetoond in deze aandoeningen waarbij een mediane tijd van 3.5 minuten nodig was om deze uitkomstmaat compleet in te vullen.

In **hoofdstuk 9** wordt een studie beschreven waarbij voor het eerst de invloed van vermoeidheid onderzocht werd op het functioneren en de kwaliteit van het leven in patiënten met immuun-gemedieerde polyneuropathieën. Ernstige vermoeidheid werd gevonden in een hoog percentage (80%) van deze patiënten en bleek een ernstig invaliderend symptoom te zijn. Vermoeidheid bleek op zichzelf te staan en geen relatie te hebben met zwakte, gevoelsstoornissen of duur van de ziekte. Bij patiënten met GBS bleek vermoeidheid gerelateerd te zijn aan psychosociale domeinen van de kwaliteit van leven schaal, de SF-36. In tegenstelling tot deze bevindingen bleek vermoeidheid bij patiënten met een chronische polyneuropathie (CIDP of MGUSP) eerder een relatie te hebben met de meer fysiek georiënteerde domeinen van de SF-36. Vanuit deze bevindingen werd verondersteld dat GBS patiënten de klap van het snel ontwikkelen van neurologische verschijnselen waarschijnlijk moeizaam te boven komen, terwijl de patiënten met chronische verschijnselen meer gebukt gaan onder de dreiging van verdere fysieke achteruitgang, waardoor zij meer geneigd zouden zijn om klachten, zoals vermoeidheid, te relateren aan het fysiek lijden.
Voor de evaluatie van vermoeidheid is gebruik gemaakt van een schaal, de zogenaamde: "Fatigue severity scale". De simpelheid, validiteit, en betrouwbaarheid van deze uitkomstmaat werden ook in deze patiënten gedemonstreerd.

In **hoofstuk 10** wordt een studie beschreven waarin de kwaliteit van leven wordt onderzocht bij patiënten met immuun-gemedieerde polyneuropathieën. De kwaliteit van leven werd

geregistreerd met behulp van de SF-36 kwaliteit van leven schaal. De verkregen gemiddelde waarden voor de 8 SF-36 domeinen alsmede de waarden voor de totale fysieke en mentale samengestelde eind-onderdelen werden vergeleken met de gerapporteerde gemiddelde waarden die recentelijk gepubliceerd zijn voor een groep van 1742 gezonde Nederlandse inwoners. De SF-36 beschreef op een adequate manier de algemene tekortkomingen in de kwaliteit van het bestaan in deze patiënten waardoor een meer compleet beeld verkregen werd van de grootte en uitgebreidheid van de inslag van deze aandoeningen op iemands bestaan. Bij de groep patiënten die in de tijd gevolgd werd, werd een geleidelijk herstel op alle gebieden gevonden, waarbij het mentale gedeelte sneller dan het fysieke de normaalwaarden grens bereikte. Gesuggereerd werd dat een intensieve begeleiding van deze patiënten meest waarschijnlijk als een belangrijke bijdrage gezien moest worden bij het algeheel herstellen van deze patiënten, naast uiteraard de medicijnen die gegeven zijn. Voorts werden acceptabele klinimetrische waarden verkregen voor de SF-36 in het geheel. Op basis hiervan werd geconcludeerd dat deze schaal als adjunct toegepast kon worden bij patiënten met met immuun-gemedieerde polyneuropathieën.

Van alle klinimetrische vereisten is responsiviteit het minst onderzocht in de geneeskunde. In **hoofdstuk 11** wordt een studie beschreven waarin, naast validiteit en betrouwbaarheid, verschillende statistische technieken om 'responsiviteit' te meten vergeleken worden. Twee nieuwe, zogenaamde 'verdelingsvrije' responsiviteit technieken (lees: technieken die geen rekening houden met het wel of niet normaal verdeeld zijn van een gemeten kwaliteit), werden hierbij geintroduceerd. De vergelijking tussen de diverse technieken werd verricht door gebruik te maken van de geselecteerde impairment en disability schalen. In het algemeen rangschikten de gebruikte methoden om responsiviteit te meten de schalen steeds op dezelfde volgorde, waardoor de dokter gemakkelijk kon kiezen tussen schalen die dezelfde validiteit en betrouwbaarheid bleken te hebben. Van de onderzochte schalen bleken de "overall disability sumscore", die informatie verschaft over de functionaliteit van de armen/handen en benen, the MRC sumscore en de Vigorimeter steeds gerangschikt te zijn tussen de schalen met de hoogste waarden voor responsiviteit. Voorts bleken deze schalen onderling goede correlaties te vertonen. Derhalve werd het gebruik van deze schalen primair gesuggereerd bij de evaluatie van uitkomsten op het niveau van impairment en disability bij patiënten met immuun-gemedieerde polyneuropathieën.

In **hoofdstuk 12** worden de associaties tussen geselecteerde impairment, disability en handicap schalen beschreven bij patiënten met immuun-gemedieeerde polyneuropathieën. Deze studie vormt een ondersteuning voor het concept van de 'international classification of impairments, disabilities, and handicaps', waarin een correlatie tussen de diverse niveau's gesuggereerd wordt. Impairment schalen hebben tweederde van disability verklaard en circa de helft van handicap. De MRC sumscore en de knijpkracht (Vigorimeter) hebben de grootste bijdrage geleverd in deze modellen. Disability schalen hebben 76% van handicap verklaard waarbij de overall disability sumscore de belangrijkste verklarende variabele was. De combinatie van impairment en disability schalen heeft 77% van de gemeten handicap stoornissen verklaard. Echter, omdat de gevonden associaties niet absoluut waren worden mogelijke andere verklarende variabelen, zoals pijn, psychologische factoren als angst, depressie, coping mechanisme, motivatie, sociale ondersteuning en lichamelijke conditie, aangedragen voor toekomstige studies.

Uiteindelijk wordt in **hoofdstuk 13** naast een samenvatting in de Engelse taal, suggesties aangedragen voor toekomstige studies op de diverse terreinen.

Dankwoord

Het is nogal cliché om te zeggen dat je niet in je eentje een proefschrift schrijft, maar het is wel de werkelijkheid. Als ik bij de verdediging sta, dan verdedig ik niet alleen mijn gedachten die verwerkt zijn in de diverse hoofdstukken, maar alle verwerkte ondersteuningen van vele anderen. Door dit te beseffen kan ik gerust zeggen dat ik een enorme steun in de rug ervaar, zelfs bij dit laatste gebeuren. Mijn dank voor al deze mensen is immens groot dat woorden te kort schieten om dit te beschrijven.
Allereerst wil ik lieve God bedanken voor de inspiratie en het doorzettingsvermogen die hij mij gegeven heeft. Alleen hij kan weten hoe sterk ik aan mijzelf getwijfeld heb gedurende de turbulente periodes van deze studie. Moeilijk was het als ik vanuit mijn eigen optiek kijk. Echter, door Gods belichte wegen te volgen konden vele op het oog niet overbrugbare barricades met redelijk gemak overwonnen worden – de ingeving om de vereniging van GBS patiënten te bellen toen wij niet voldoende patiënten bleken te hebben – 's nachts wakker worden en precies weten wat je schrijven moet – in een richting naar een oplossing zoeken en pas achteraf beseffen dat deze de juiste richting was – zijn slechts enkele voorbeelden van begeleiding en ondersteuning van de Heer.
Mijn lieve echtgenote en dochter die met veel geduld en relatief weinig tegensputteren mijn regelmatig verzuim thuis wisten op te vangen en altijd voor mij klaar stonden, wil ik bedanken. Mijn lieve Els, die praktisch alle artikelen (inclusief mijn dankwoord) gecontroleerd heeft op grammaticale misconstructies – je weet dat ik dit nooit had kunnen bewerkstelligen als jij niet in mij geloofde en de nodige steun gaf om mij staande te houden. Lieve Els, al hetgeen ik als mens bereikt heb komt door jou en doe ik voor jou. Ik wil je bedanken voor dit alles en hoop met de zegening van God dat wij nog vele jaren in volle gezondheid van elkaar mogen genieten. Een mooier en krachtiger mens als partner kan iemand niet wensen. Bedankt lieverd!
Mijn lieve dochter Jamie Alarice Kim: deel van mijn tranen, deel van mijn kloppend hart, mooi om te zien hoe jij met projectjes en dergelijke ook naast mij opgroeit. Bedankt dat jij mijn leven ook zo verrijkt; God zegen ook jou wegen lieverd en kijk maar wat je later wilt worden ..... papa vind alles best, als je maar gelukkig bent.
Mijn lieve ouders, die in vrede mogen rusten: Lieve papai i Mabea, Dios kwida bosonan i mi pensamentu over di bosonan su presencia; danki pa tur kos ku bosonan a ofresemi i pa e sosten grandi; ku boso aporte tambe mi por bisa ku mi por a logra algu; masha danki pa tur kos; mi sa ku bosonan lo ta presente tambe.
Aan mijn broers en zusters: Ook bedankt voor jullie ondersteuning en het feit dat jullie ook in mij geloofden; bij het opsommen van al dit, besef ik hoe rijkelijk ik bedeeld ben met zoveel broers en zusters; door het verlies van mijn ouders besef ik ook dat onze banden duidelijker versterkt zijn; ik hoop dat dit voor altijd mag blijven bestaan en dat wij in volle glorie elkaar mogen steunen in alles.

Prof. dr. F.G.A. van der Meché, beste promotor of beste Frans, bedankt voor het vertrouwen en ondersteuning die ik van jou in de afgelopen jaren heb gekregen. Dat je mij aannam als assistent in opleiding zonder veel over mij te weten, de mogelijkheid heb gegeven om mij te ontplooien tot neuroloog, het feit dat je mij ook geleerd heb onderzoek te verrichten, voor dit alles wil ik je bedanken. Ondanks het feit dat dit onderzoek als een soort "politieke zet" is begonnen, ben ik blij dat het geheel onder jou auspiciën zo is uitgepakt. Ik dank je ook Frans

voor onze persoonlijke gesprekken en je persoonlijke belangstelling in de moeilijke tijden die ik had.
Dr. P.A. van Doorn, beste Pieter of te wel mijn klinisch georiënteerde co-promoter: Bedankt dat je mij niet alleen in mijn klinische blik aangescherpt hebt, mijn neurologisch denken en handelen verrijkt hebt, maar ook voor het feit dat je veel structuur gegeven hebt aan mijn wetenschappelijke vorming. Ik ben je zeer erkentelijk. Al onze overleggingen en afwegingen, het geduld dat je moest opbrengen om mij in zo'n taaie materie als meetkunde te blijven stimuleren, heb ik altijd zeer op prijs gesteld. 'The fine-tuning and final touch' van al het geschreven kon jij beter dan wie dan ook. Velen van ons weten misschien niet wat Cornelis Poortvliet zou zeggen over jou: ..... wij weten het in ieder geval wel! Ik denk persoonlijk dat het niet onterecht is. Pieter, bedankt voor jou enorme steun en ik hoop dat wij in de toekomst nog iets op wetenschappelijk gebied gezamenlijk kunnen blijven doen.

Dr. Ir. P.I.M. Schmitz, beste vriend Paul of te wel mijn statistisch georiënteerde co-promotor: Gezamenlijk met je klinisch georiënteerde collega co-promotor wisten jullie iets moois van dit onderzoek te bewerkstelligen. Klinische ideëen hebben en het verzamelen van gegevens is slechts de basis tot de vorming van een wetenschappelijk huis. Statistiek is in mijn ogen nog steeds een goochel-truc die voor weinigen weggelegd is. Je enthousiasme, je vastberadenheid, je creativiteit (lees: gegoochel) om het geheel tot een geordend geheel om te turnen, zijn slechts enkele aspecten die jouw persoon kenmerken. We hebben vele uren lachend, serieus filosoferend doorgebracht en gelukkig keihard doorgewerkt. Het resultaat is nu zichtbaar denk ik. Paul, bedankt voor alles; ik had dit niet kunnen realiseren als je niet zo vastberaden was. Pas goed op jezelf en nogmaals danki.
Prof. Toyka wil ik bedanken voor zijn advies om de Rydel-Seiffer stemvork te gebruiken. Prof. Passchier wil ik bedanken voor zijn advies om de SF-36 kwaliteit van leven schaal te gebruiken. Prof. Hughes bedank ik ook voor de bijdrage die hij heeft geleverd bij het construeren van de diverse schalen.
De participanten in deze studie wil ik ook bedanken voor hun ondersteuning. In het bijzonder alle gezonde vrijwilligers uit het AZR, uit de gemeente Haarlemmermeer, met name uit Nieuw-Vennep, Hoofddorp en Leimuiden, bedankt voor jullie steun. Ook de onderzoekers dank ik voor hun bijdrage aan het geheel. Ron Meijer, Frank Opstelten, Johnny Samijn, Wibe Moll en Rinske van Koningsveld hebben uren op de polikliniek gespendeerd met het afnemen van metingen bij vele patiënten. Bedankt voor jullie steun en geduld. Speciale dank gaat uit naar Johnny Samijn, die niet alleen mijn directe onderzoekspartner vormde, maar veel geografisch onderwijs over Nederland (gedurende onze uitstapjes naar enkele patiënten) aan mij wist over te dragen. Speciale dank ook voor Rinske voor het starten met het vibratie onderzoek in een periode die voor mij te belastend was. Jou vastberadenheid werkte als stimulans voor mij.

De vereniging spierziekten Nederland, onderdeel GBS patiënten vereniging, wil ik bedanken voor hun ondersteuning. Zonder deze steun konden wij niet zoveel patiënten recruteren. Speciale dank gaat uit naar alle patiënten die een bijdrage hebben geleverd in deze studie. Sommige van jullie waren zelfs uren onderweg om bij ons te zijn. Mede door jullie bijdrage is het geheel gelukt en zijn wij meerdere keren op het juiste spoor gezet om een bepaalde onderwerp, bijvoorbeeld vermoeidheid, uit te zoeken. Ik hoop dat dit geheel een bijdrage zal leveren tot verdere verheldering van jullie lijden.

Naast Johnny, wil ik in het bijzonder mijn studie-maatje Stef Bakker bedanken voor zijn steun in vele opzichten. Niet alleen als paranimf, maar ook door de jaren heen wist jij Stef mij gaande te houden door jou steun als collega en vriend. Ik hoop dat jij ook binnenkort je thesis kan afronden en dank je nogmaals voor alle steun.
Alle stafleden neurologie van het AZR en oude collega assistenten wil ik bedanken voor hun steun en bijdrage aan mijn opleiding. De verpleging neurologie van het AZR, de secretaressen en doktersassistenten wil ik ook bedanken voor hun geduld en ondersteuning gedurende mijn opleiding. In het bijzonder wil ik bedanken Lydia Loman, onze voormalige octopus-secretaresse, voor haar steun in het begin van mijn wetenschappelijke carrière. Lydia, wij zijn jou niet vergeten en ik wens je persoonlijk heel veel geluk.

Laurens van Briemen, Ton Mus, Chris Sieradzan, Djo Hasan en Etienne Samson (Algemeen Dagblad Carribische editie) dank ik voor hun ondersteuning in allerlei technische aspekten en computer programmatuur. Speciale dank gaat uit naar uitgeverij van Groenigen – Witte Weekblad' te Nieuw-Vennep, die kosteloos een oproep hebben geplaatst voor de knijpkracht studie.
Mijn directe collega's, Roberto Rico en Wim van der Kamp, alsmede onze secretaressen en ENF-medewerksters, dank ik voor hun steun in de periode van afronding van mijn thesis.
Jacob (Ado) Gravenhorst, Manrique Lodewijks en Earl Esseboom dank ik persoonlijk voor hun steun als ware vrienden door de jaren heen. Alle kennissen die op een of andere wijze een bijdrage hebben geleverd aan het realiseren van deze studie en die niet expliciet vermeld zijn, dank ik middels dit schrijven.

List of abbreviations

ADS	= arm disability scale
AIDP	= acute inflammatory demyelinating polyneuropathy
Anova	= analysis-of-variance
BP	= body pain
CHART	= Craig handicap assessment and reporting technique
CIDP	= chronic inflammatory demyelinating polyneuropathy
CIQ	= community integration questionnaire
CMDN	= chronic multifocal demyelinating neuropathy
DM	= diabetes mellitus
ES	= effect size
ESS	= environmental status scale
FP	= female patients
F-score	= Hughes' functional grading scale
FSS	= fatigue severity scale
GBS	= Guillain-Barré syndrome
GHP	= general health perception
GS	= grip strength
HAS	= handicap assessment scale
HC	= healthy controls
HF	= healthy females
HM	= healthy males
ICIDH	= international classification of impairments, disabilities, and handicaps
ICIDH-2	= international classification of impairments, activities, and participation
INCAT	= inflammatory neuropathy cause and treatment group
IQR	= inter-quartile range
IQRchange	= 75^{th} percentile – 25^{th} percentile
ISS	= Incat sensory sumscore
IVIg	= intravenous immunoglobulin
kPa	= kiloPascal
LDS	= leg disability scale
LH	= left hand
MCS	= mental component summary score
MGUS(P)	= monoclonal gammopathy of undetermined significance (associated polyneuropathy)
MH	= mental health
MMN	= multifocal motor neuropathy
MP	= male patients
MRC	= medical research council
MS	= multiple sclerosis
9HPT	= nine hole peg test
ODSS	= overall disability sumscore
PCS	= physical component summary score
PhF	= physical functioning
PNP	= polyneuropathy

RA	= rheumatoid arthritis
RFE	= role functioning – emotional
RFPh	= role functioning – physical
RH	= right hand
RIHS9	= Rotterdam 9 items handicap scale
RS	= Rydel-Seiffer
SAIDP	= subacute inflammatory demyelinating polyneuropathy
SD	= standard deviation
SDres	= residual standard deviation around the mean regression line
SF	= social functioning
SF-36	= medical outcome study 36-items short form health survey
SFFS	= short form fatigue scale
SIP	= sickness impact profile
SLE	= systemic lupus erythematosus
Srcc	= Spearman's rank correlation coefficient
SRM	= standardised response mean
ss	= sumscore
10MW(P)T	= ten meters walking (performance) test
VAS	= visual analogue scale
Vit	= vitality
VM	= Vigorimeter
WHO	= World health organisation

Curriculum vitae

Name:	**Ingemar Sergio José Merkies** (formerly known as Martina)
Born:	**September 9, 1964, Curaçao, Netherlands Antilles**
1984:	Graduated from Grammar School, Amsterdam, The Netherlands
1985-1987:	Spanish, Free University Hospital Amsterdam, Amsterdam, The Netherlands
1984-1991:	Medical School, Free University Hospital Amsterdam, Amsterdam, The Netherlands (cum laude)
1991-1992:	Resident in Neurology (Agnio), St. Elisabeth Hospital, Haarlem, The Netherlands
1992:	Resident in Neurology (Agnio), Slotervaart Hospital, Amsterdam, The Netherlands
1993-1994:	Resident in Neurology (Agnio), St. Elisabeth Hospital, Willemstad, Curaçao, Netherlands Antilles
1994-2000:	Resident in Neurology (Agio), Dept. of Neurology, Erasmus Medical Centre Rotterdam, Rotterdam, The Netherlands
Currently:	Stafmember department of Neurology, St. Elisabeth hospitaal, Willemstad, Curaçao, Netherlands Antilles.

List of publications

1 **Martina ISJ**, van Doorn PA. Spinal cord compression due to extramedullary haematopoiesis in thalassaemia: a case report and review of the literature. *J Neurology 1996;243:364-366.*

2 Kamphuis DJ, **Martina ISJ**, Keli SO, Balentien EGD, Top K. Seroprevalentie van humaan T-cellymfotroop virus type I op Curaçao en voorspellende variabelen voor die infectie. *Ned Tijdschr Geneeskd 1998;142:89-92.*

3 **Martina ISJ**, van Koningsveld R, van Doorn PA, Schmitz PIM, van der Meché FGA, for the Inflammatory Neuropathy Cause and Treatment (*INCAT*) group. Measuring vibration threshold with a graduated tuning fork in normal ageing and in patients with polyneuropathy. *J Neurol Neurosurg Psychiatry 1998;65:743-747.*

4 **Martina ISJ**, van Doorn PA, Schmitz PIM, Meulstee J, van der Meché FGA. Chronic motor neuropathies: response to interferon-β1a after failure of conventional therapies. *J Neurol Neurosurg Psychiatry 1999;66:197-201.*

5 **Merkies ISJ**, Schmitz PIM, Samijn JPA, van der Meché FGA, van Doorn PA, for the Inflammatory Neuropathy Cause and Treatment (*INCAT*) group. Fatigue in immune-mediated polyneuropathies. *Neurology 1999;53:1648-1654.*

6 **Merkies ISJ**, Schmitz PIM, van der Meché FGA, van Doorn PA, for the Inflammatory Neuropathy Cause and Treatment (*INCAT*) group. Psychometric evaluation of a new sensory scale in immune-mediated polyneuropathies. *Neurology 2000;54:943-947.*

7 **Merkies ISJ**, Schmitz PIM, van der Meché FGA, van Doorn PA, for the Inflammatory Neuropathy Cause and Treatment (*INCAT*) group. Reliability and responsiveness of a graduated tuning fork in immune-mediated polyneuropathies. *J Neurol Neurosurg Psychiatry 2000;68:669-671.*

8 **Merkies ISJ**, Schmitz PIM, Samijn JPA, Toyka KV, van der Meché FGA, van Doorn PA, for the Inflammatory Neuropathy Cause and Treatment (*INCAT*) group. Assessing grip strength in healthy individuals and patients with immune-mediated polyneuropathies. *Muscle & Nerve 2000;23:1393-1401.*

9 Van de Berg-Vos RM, Van den Berg LH, Franssen H, van Doorn PA, **Merkies ISJ**, Wokke JHJ. Treatment of multifocal motor neuropathy with interferon-β1a. *Neurology 2000;54:1518-1521.*

10 Ang CW, van Doorn PA, Endtz HPh, **Merkies ISJ**, Jacobs BC, de Klerk MA, van Koningsveld R, van der Meché FGA. A single case of Guillain-Barré syndrome in a family with campylobacter jejuni enteritis. *J neuroimmunol 2000;111:229-233.*

11 **Merkies ISJ**, Schmitz PIM, van der Meché FGA, Samijn JPA, van Doorn PA, for the Inflammatory Neuropathy Cause and Treatment (*INCAT*) group. Quality of life complements traditional outcome measures in immune-mediated polyneuropathies. *Submitted*

12 **Merkies ISJ**, Schmitz PIM, van Doorn PA, Samijn JPA, van der Meché FGA, for the Inflammatory Neuropathy Cause and Treatment (*INCAT*) group. Validity, Reliability, and comparative responsiveness of impairment and disability scales in immune-mediated polyneuropathies. *Submitted*

13 **Merkies ISJ**, Schmitz PIM, van der Meché FGA, Samijn JPA, van Doorn PA, for the Inflammatory Neuropathy Cause and Treatment (*INCAT*) group. Psychometric evaluation of an overall disability scale in immune-mediated polyneuropathies. *Submitted.*

14 **Merkies ISJ**, Schmitz PIM, van der Meché FGA, Samijn JPA, van Doorn PA, for the Inflammatory Neuropathy Cause and Treatment (*INCAT*) group. The development and psychometric evaluation of a new handicap scale in immune-mediated polyneuropathies. *Submitted.*

15 **Merkies ISJ**, Schmitz PIM, van der Meché FGA, Samijn JPA, van Doorn PA, for the Inflammatory Neuropathy Cause and Treatment (*INCAT*) group. Connecting impairment, disability, and handicap in immune-mediated polyneuropathies. *Submitted.*

Appendix I

Impairment scales

Medical Research Council sumscore

Range: 0 (total paralysis) to 60 (normal strength)

Muscle strength was assessed of six muscle groups (arm abductors, forearm flexors, wrist extensors, hip flexors, knee extensors, foot dorsal flexors) at both sides. The MRC scale was used to score each muscle group and the scores are given in full numbers only. For each muscle group a standardised joint/limb position as well as the point at which counter-force is administered was pre-defined and taken when assessing muscle strength.

MRC grades

0 = no movement
1 = palpable contraction, but no visible movement
2 = movement but only with gravity eliminated
3 = movement against gravity (more or less full range)
4 = movement against resistance, but weaker than normal
5 = normal power

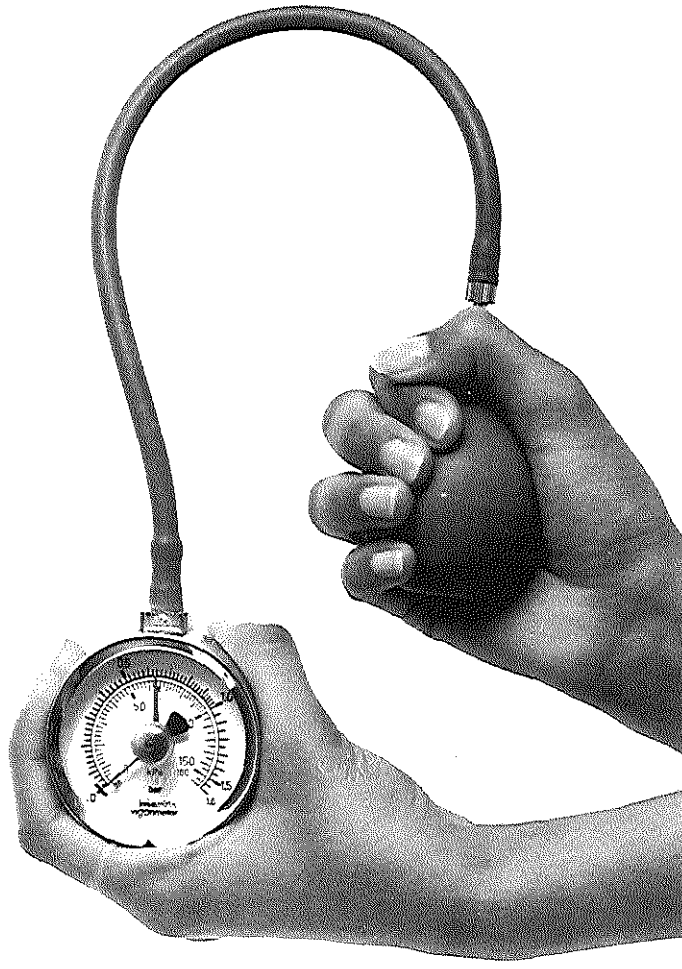
Definition of joint/limb position and point of counter-force for the muscles of the MRC sumscore. Also, a more extensive definition of some Medical Research Grading scores for various muscle groups

Muscle group	Position of the patient: Joint/limb starting position	Counter-force point of the investigator	MRC grading	Definition
Arm abductors	Patient sitting and arms hanging alongside the body	Just proximal to elbow joint	Grade 2	Abduction between 0 - < 90° (without counter-force)
			Grade 3	90° abduction against gravity (without counter-force)
Forearm flexors	Patient sitting and upper arm supported in a horizontal plane	Patient's upper arm is supported at elbow; point of counter-force: volar site wrist joint	Grade 2	Flexion of forearm between 0 - < 90° (without counter-force)
			Grade 3	90° flexion of arm against gravity (without counter-force)
Wrist extensors	Patient sitting and forearm supported in a horizontal plane	Patient's forearm supported at the volar site of wrist; point of counter-force: dorsal part hand	Grade 2	Extension between "hand-drop position" and < 45° from horizontal plane (without counter-force)
			Grade 3	At least 45° extension from horizontal plane (without counter-force)
Hip flexors	Patient lying supine and leg fully extended	Just proximal to knee joint	Grade 2	Hip Flexion < 45° (without counter-force)
			Grade 3	Hip Flexion at least 45° (without counter-force)
Knee extensors	Patient lying supine; Hip flexion at 45°	Upper leg supported just proximal to knee joint at dorsal aspect; point of counter-force: just proximal to ankle joint	Grade 2	Extension between starting point and full extension
			Grade 3	Full extension
Foot dorsal flexors	Patient sitting	Dorsal aspect of foot	Grade 2	Dorsal flexion < 90° at ankle joint (without counter-force)
				Dorsal flexion of at least 90° at ankle joint (without counter-force)

Hand-held Martin Vigorimeter

Range: 0 (no grip strength) to 160 (kiloPascal).

The *Vigorimeter* is an instrument to measure grip strength. The pressure in the bulb is registered on a manometer via a rubber junction tube and expressed in kiloPascals (kPa). According to the manufacturer's recommendation, the small bulb was used in the ages up to 10 years and the medium sized bulb was applied in the remaining participants. Participants were examined according to the assessment recommendations by the American Society of Hand Therapists (shoulder adducted and neutrally rotated, elbow flexed at 90°, forearm in neutral position, and wrist between 0° and 30° dorsiflexion and between 0° and 15° of ulnar deviation). The bulb was positioned in the palm of the participant's hand with the air tube extending out between the individual's thumb and index finger, and with the fingers wrapped around the bulb so that the fingers touched the surface of the bulb as much as possible.

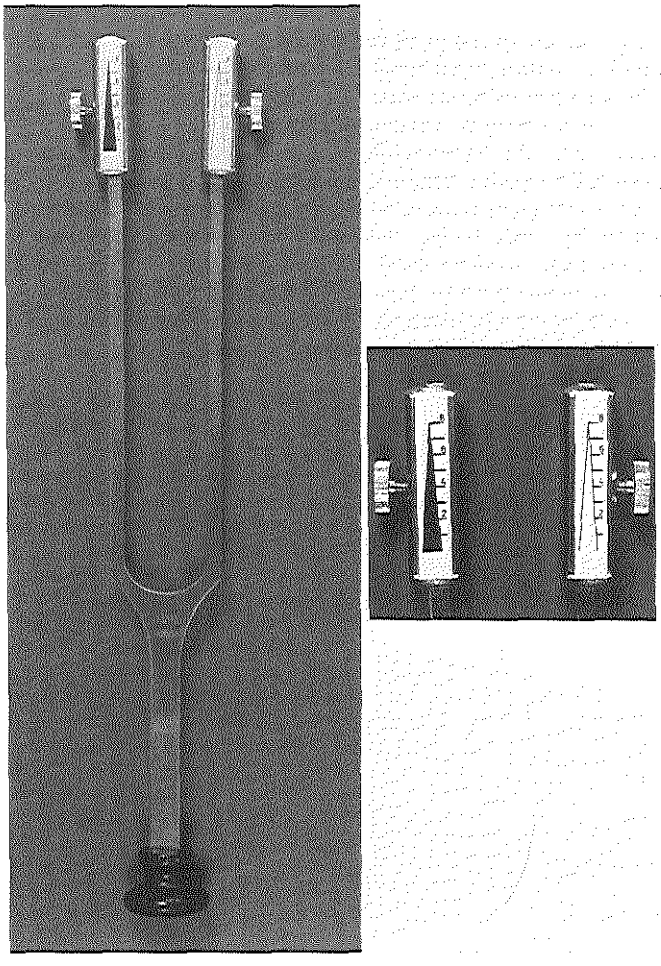


Rydel-Seiffer graduated tuning fork

Range: 0 (minimum vibration score) to 8 (maximum vibration score).

The *Rydel-Seiffer* tuning fork is a graduated fork that determines the ability of individuals to discriminate various vibration intensities. The two arms of this tuning fork bear calibrated weights at their extremities. Once the arms are swinging, the fork vibrates at 64 Hz and the triangles on the weights appear double. The intersection of these two virtual triangles moves from 0 to 8 in an exponential way with decreasing vibration-amplitude of the arms. The vibration extinction threshold is considered as the nearest value to the apparent point of intersection of the virtual triangles when the subject indicates that vibration is no longer perceived.

Sites of examination: index finger (dorsum distal interphalangeal joint), ulnar styloid process, medial humerus epicondyle, acromio-clavicular joint, hallux (dorsum interphalangeal joint), medial malleolus, patella, anterior superior iliac spine.



"INCAT" sensory sumscore

Pinprick Sensation Sites of examination + Corresponding grades		**Vibration Sensation** Sites of examination + Corresponding grades		**2-point discrimination** Site of examination + corresponding grades
Arms	Legs	Arms	Legs	Index finger[K]
0 = normal sense at index finger[A]	0 = normal sense at hallux[F]	0 = normal sense at index finger[A]	0 = normal sense at hallux[F]	0 = normal sense (≤ 4 millimetres)
Abnormal sense	**Abnormal sense**	**Abnormal sense**	**abnormal sense**	**abnormal sense**
1 = at index finger[B]	1 = at hallux[G]	1 = at index finger[B]	1 = at hallux[G]	1 = 5-9 mm
2 = at wrist[C]	2 = at ankle[H]	2 = at wrist[C]	2 = at ankle[H]	2 = 10-14 mm
3 = at elbow[D]	3 = at knee[I]	3 = at elbow[D]	3 = at knee[I]	3 = 15-19 mm
4 = at shoulder[E]	4 = at groin[J]	4 = at shoulder[E]	4 = at groin[J]	4 = 20 mm or more

Pinprick and *vibration* sense examination took place from distal to proximal and only the highest extension of dysfunction of the most affected arm and leg was recorded separately for both qualities.

Pinprick was tested using the sharp end of an esthesiometer. Patients were asked to indicate whether they experienced the pinprick as normal or abnormal. Paresthesia, dysesthesia or hyperesthesia were scored as abnormal. We seek for a normal reference point (e.g. sensation at the face), if a patient was experiencing problems indicating whether the pinprick was normal or not.

Vibration was assessed using the Rydel-Seiffer graduated tuning fork and the obtained values were compared with the published normative vibration threshold values.

ISS composition: pinprick arm grade [range: 0-4] + vibration arm grade [range: 0-4] + pinprick leg grade [range: 0-4] + vibration leg grade [range: 0-4] + 2-point discrimination grade [range: 0-4]. *Sites of examination*: [A & B]=index finger (dorsum distal interphalangeal joint); [C]=ulnar styloid process; [D]=medial humerus epicondyle; [E]=acromio-clavicular joint; [F & G]=hallux (dorsum interphalangeal joint); [H]=medial malleolus; [I]=patella; [J]=anterior superior iliac spine. [K]=index finger (ventral side; distal phalanx).

ISS Range: 0 ("no sensory deficit") to 20 ("most severe sensory deficit").

Fatigue severity scale
Range: 0 (no signs of fatigue) to 7 (most disabling fatigue score).

English version

Fatigue severity scale

	1=strongly disagree; 2=mainly disagree; 3=partially disagree; 4=do not agree/disagree; 5=partially agree; 6=mainly agree 7=strongly agree (circle one answer per question)							
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	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

Patients are instructed to choose a number from 1 to 7 that indicates their degree of agreement with each statement (1=strongly disagree; 7=strongly agree). The mean of these 9 questions is considered 'the fatigue score' for each patient.

Dutch version

Fatigue severity scale

	1=volledig oneens; 2=grotendeels oneens; 3=gedeeltelijk oneens; 4=niet oneens/niet eens; 5=gedeeltelijk eens; 6=grotendeels eens 7=volledig eens (omcirkel één keuze per vraag)							
1	Ik ben minder gemotiveerd om dingen te doen als ik moe ben	1	2	3	4	5	6	7
2	lichamelijke inspanning leidt tot vermoeidheid	1	2	3	4	5	6	7
3	ik ben snel moe/vermoeid	1	2	3	4	5	6	7
4	moeheid/vermoeidheid belemmert me in mijn lichamelijk functioneren	1	2	3	4	5	6	7
5	moeheid/vermoeidheid leidt voor mij vaak tot problemen	1	2	3	4	5	6	7
6	moeheid/vermoeidheid verhindert langdurige lichamelijke inspanning	1	2	3	4	5	6	7
7	moeheid/vermoeidheid beinvloedt de uitvoering van bepaalde taken en verplichtingen	1	2	3	4	5	6	7
8	moeheid/vermoeidheid behoort tot mijn drie voornaamste klachten	1	2	3	4	5	6	7
9	Moeheid/vermoeidheid beinvloedt mijn werk, gezinsleven of sociale activiteiten	1	2	3	4	5	6	7

Patiënten worden geinstrueerd om bij elke vraag 1 antwoord te omcirkelen dat het beste past bij hun situatie. Het gemiddelde van deze 9 vragen wordt berekend en beschouwd als 'the fatigue score'.

Appendix II

Disability scales

Functional grading scale (f-score; Hughes' disability scale)
The *functional grading scale* (*f-score*) assesses the functional ability of the patients.

Functional grading scale
0 = normal; no symptoms or signs
1 = minor neurological symptoms or signs and able to run
2 = able to walk at least 10 meters, but unable to run
3 = able to walk 10 meters with a walker or support
4 = bedridden or chair bound (unable to walk 10 meters with a walker or support)
5 = requiring artificial ventilation for at least part of the day
6 = dead

Overall disability sumscore

Arm disability scale – Function checklist	**Not affected**	**Affected but not prevented**	**Prevented**
Dressing upper part of body (excluding buttons/zips)	O	O	O
Washing and brushing hair	O	O	O
Turning a key in a lock	O	O	O
Using knife and fork (spoon: is applicable if the patient never uses knife and fork)	O	O	O
Doing/undoing buttons and zips	O	O	O
Arm grade			
0 = Normal			
1 = minor symptoms or signs in one or both arms but not affecting any of the functions listed			
2 = moderate symptoms or signs in one or both arms affecting but not preventing any of the functions listed			
3 = severe symptoms or signs in one or both arms preventing at least one but not all functions listed			
4 = severe symptoms or signs in both arms preventing all functions listed but some purposeful movements still possible			
5 = severe symptoms and signs in both arms preventing all purposeful movements			
Leg disability scale – Function checklist	**No**	**Yes**	**Not applicable**
Do you have any problem with your walking	O	O	O
Do you use a walking aid	O	O	O
How do you do usually get around for about 10 meters			
Without aid	O	O	O
With one stick or crutch or holding to someone's arm	O	O	O
With two sticks or crutches or one stick or crutch and holding to someone's arm	O	O	O
With a wheelchair	O	O	O
If you use a wheelchair: can you stand and walk a few steps with help	O	O	O
If you are restricted to bed most of the time, are you able to make some purposeful movements	O	O	O
Leg grade			
0 = Walking is not affected			
1 = walking is affected but does not look abnormal			
2 = walks independently but gait looks abnormal			
3 = usually uses unilateral support to walk 10 meters (stick, single crutch, one arm - 25 yards)			
4 = usually uses bilateral support to walk 10 meters (sticks, crutches, two arms - 25 yards)			
5 = usually uses wheelchair to travel 10 meters (25 yards)			
6 = restricted to wheelchair, unable to stand and walk few steps with help but able to make some purposeful leg movements			
7 = restricted to wheelchair or bed most of the day, preventing all purposeful movements of the legs (e.g. unable to reposition the legs in bed)			

Overall disability sumscore = Arm disability scale (range: 0-5) + Leg disability scale (range: 0-7)
Range: 0 (no signs of disability) to 12 (maximum disability).
For the arm disability scale: Allocate one arm grade only by completing the Function checklist. Indicate whether each function is 'affected', 'affected but not prevented' or 'prevented'. For the leg disability scale: Allocate one leg grade only by completing the Functional questions.

Ten-metres walking test

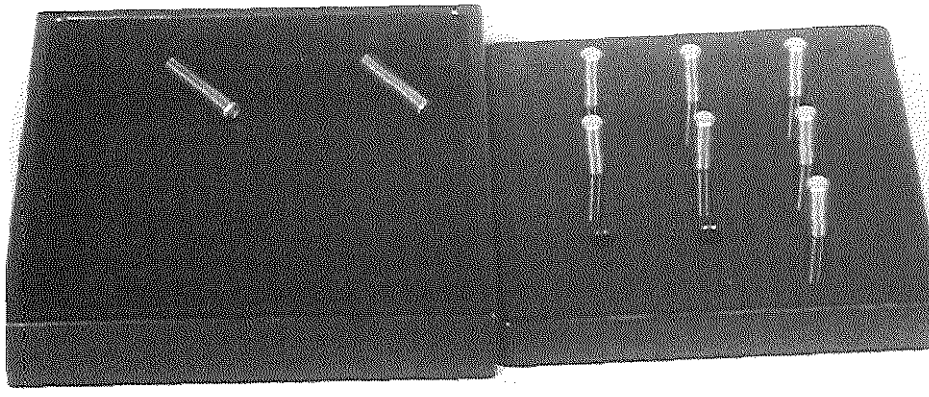
The patient is asked to walk at his or her own preferred speed, using whatever aid needed [including personal support if wanted]. They are then asked to walk 10 metres in a straight line. The patient is timed over the distance in seconds. The result can be reported as the number of seconds taken, or as speed [metres/second]. Any aid or assistance should be recorded. [Cave: ceiling effect once the patient reached normal walking speed; it is possible then to consider measuring running speed or endurance].

Nine-hole peg test

Equipment

9 dowels; 9 mm diameter, 32 mm long. base with 9 holes [10 mm diam, 15 mm deep] spaced 15 mm apart in three rows of three holes. Lid to base, with tray 100 mm square and 10 mm deep to hold pegs.

Instructions: A brief interview preceded the test to determine hand dominance. If the patient reported to use both hands equally, the hand used to write will be considered as the 'dominant' hand. Both hands were tested, starting with the dominant hand. The patient had to sit at table [in a chair or in bed], and was asked to place pegs in holes. The observer timed from start to end, but the stopwatch could also be stopped at 1 minute and record number of pegs placed was then noted. This instrument is also useful for measuring the disabling effects of sensory loss and ataxia. Most people complete the test in 18 seconds.



The Modified Rankin Scale

0 = no symptoms at all
1 = no significant disability despite symptoms: able to carry out all usual duties and activities
2 = slight disability: unable to carry out all previous activities but able to look after own affairs without assistance
3 = moderate disability: requiring some help, but able to walk without assistance
4 = moderately severe disability: unable to walk without assistance or unable to attend to own bodily needs without assistance
5 = severe disability: bedridden, incontinent, and requiring constant nursing care and attention

Appendix III

Handicap scale

Rotterdam nine items handicap scale

English version

1. Mobility indoors
Are you able to move from room to room, negotiating doors, carpets and polished surfaces?
0 = not applicable
1 = unable to move between rooms
2 = moves between rooms mostly with help of another person
3 = moves between rooms most of the time independent; sometimes needing help of another person
4 = moves between rooms totally independent

2. Mobility outdoors
Are you able to move outdoors from one place to another, negotiating kerbs and uneven grounds?
0 = not applicable
1 = unable to move outdoors
2 = moves outdoors mostly with help of another person
3 = moves outdoors most of the time independent; sometimes needing help of another person
4 = moves outdoors totally independent

3. Kitchen tasks
Are you able to fulfil tasks like making a pot of tea/coffee, and serving it; are you able to collect items from a high and low cupboard, refrigerator, etcetera? (other kitchen tasks are also applicable).
0 = not applicable
1 = unable to fulfil any kitchen task
2 = able to fulfil only a minimum of these tasks; mostly needing help of another person
3 = able to fulfil the vast majority of these tasks independently; sometimes needing help of another person
4 = able to fulfil all kitchen tasks independently

4. Domestic tasks (indoors)
Are you able to fulfil house-cleaning tasks, such as vacuum cleaning, dishwashing, doing the laundry, dusting, etcetera?
0 = not applicable
1 = unable to fulfil any domestic tasks indoors
2 = able to fulfil only a minimum of these tasks; mostly needing help of another person
3 = able to fulfil the vast majority of these tasks independently; sometimes needing help of another person
4 = able to fulfil all domestic tasks indoors independently

5. Domestic tasks (outdoors)
Are you able to do the shopping, managing the garden, cleaning the car, etcetera?
0 = not applicable
1 = unable to fulfil any domestic tasks outdoors
2 = able to fulfil only a minimum of these tasks; mostly needing help of another person
3 = able to fulfil the vast majority of these tasks independently; sometimes needing help of another person
4 = able to fulfil all domestic tasks outdoors independently

6. Leisure activities (indoors)
Are you able to read a newspaper/magazine or a book, use the telephone, fulfil a hobby (other then sporting)?
0 = not applicable
1 = unable to fulfil these activities
2 = able to fulfil only a minimum of these activities; mostly needing help of another person
3 = able to fulfil the vast majority of these activities independently; sometimes needing help of another person
4 = able to fulfil all these activities independently

7. Leisure activities (outdoors)
Are you able to go to a party, theatre, movies, concerts, museums, meetings, participate in sport?
0 = not applicable
1 = unable to fulfil these activities
2 = able to fulfil only a minimum of these activities; mostly needing help of another person
3 = able to fulfil the vast majority of these activities independently; sometimes needing help of another person
4 = able to fulfil all these activities independently

8. Able to drive a car/go by bus/ride a bicycle
Are you able to drive a car, go on a bus/subway, or ride a bicycle?
0 = not applicable
1 = unable to fulfil any of these tasks
2 = able to fulfil only one of these tasks (if needed with help of another person)
3 = able to fulfil two of these tasks (if needed with help of another person)
4 = able to fulfil all these tasks independently

9. Work/study
Are you able to fulfil your prior (before becoming ill) job/study?
0 = not applicable
1 = unable to fulfil prior job/study
2 = able to fulfil (partly) adapted job/study
3 = able to fulfil partly the prior job/study
4 = able to fulfil completely prior job/study

Score-range, guidelines, and formulas for transformation of raw scores to final scores for Rotterdam 9 items handicap scale
The Rotterdam nine items handicap scale score: summation of all applicable items x 9/(9-number of not applicable items); score-range: 9 ("unable to fulfil any task/activity") to 36 ("able to fulfil all tasks/activities").

Regarding items 1 and 2: Moving from room to room or outdoors does not necessarily mean that a patient has the ability to walk. As an example: A patient can also move from one place to another in a wheelchair.

Regarding item 8: For example, if a patient does not have a driving license, it was proposed to consider this part of the question as "being fulfilled", unless it was clear that this would be absolutely impossible due to him/her illness.

Dutch version

1. Mobiliteit binnenshuis

Kunt u zich verplaatsen van kamer naar kamer, waarbij deuren, drempels, tapijten en/of gladde vloeren overbrugd worden?

0 = niet van toepassing
1 = ik kan mij niet verplaatsen van kamer naar kamer
2 = ik verplaats mij meestal met de hulp van iemand anders van kamer naar kamer
3 = ik verplaats mij grotendeels zelfstandig van kamer naar kamer
4 = ik verplaats mij volledig zelfstandig van kamer naar kamer

2. Mobiliteit buitenshuis

Kunt u zich buiten verplaatsen waarbij eventueel trottoirs en oneffenheden overbrugd worden?

0 = niet van toepassing
1 = ik kan mij niet buiten verplaatsen
2 = ik verplaats mij meestal met de hulp van iemand anders buiten
3 = ik verplaats mij grotendeels zelfstandig buiten
4 = ik verplaats mij volledig zelfstandig buiten

3. Keuken taken

Kunt u keuken taken als bijvoorbeeld een pot thee/koffie zetten en serveren; bent u hierbij in staat om voorwerpen uit kasten/lades e.d. te pakken? (andere keuken-activiteiten komen ook in aanmerking).

0 = niet van toepassing
1 = ik kan geen keuken taken uitvoeren
2 = ik voer slechts een deel van de keuken taken uit; meestal heb ik de hulp van iemand anders nodig
3 = ik voer de meeste keuken taken (grotendeels) zelfstandig uit
4 = ik voer alle keuken taken volledig zelfstandig uit

4. Huishoudelijke taken [binnenshuis]

Kunt u schoonmaak-taken zoals stofzuigen, afwassen, afstoffen, de was doen, etc.?

0 = niet van toepassing
1 = ik kan deze taken niet uitvoeren
2 = ik voer slechts een deel van deze taken uit; meestal heb ik de hulp van iemand anders nodig
3 = ik voer de meeste van deze taken (grotendeels) zelfstandig uit
4 = ik voer al deze taken volledig zelfstandig uit

5. Huishoudelijke taken [buitenshuis]

Kunt u boodschappen doen, de tuin bewerken, ramen zemen, auto wassen etc.?

0 = niet van toepassing
1 = ik kan deze taken niet uitvoeren
2 = ik voer slechts een deel van deze taken uit; meestal heb ik de hulp van iemand anders nodig
3 = ik voer de meeste van deze taken (grotendeels) zelfstandig uit
4 = ik voer al deze taken volledig zelfstandig uit

6. Ontspanningsactiviteiten [binnenshuis]

Kunt u bijvoorbeeld krant of boek lezen, telefoneren, een hobby uitvoeren?

0 = niet van toepassing
1 = ik kan deze aktiviteiten niet uitvoeren
2 = ik voer slechts een deel van deze aktiviteiten uit; meestal heb ik de hulp van iemand anders nodig
3 = ik voer de meeste van deze aktiviteiten (grotendeels) zelfstandig uit
4 = ik voer al deze aktiviteiten volledig zelfstandig uit

7. Ontspanningsactiviteiten [buitenshuis]
Bent u in staat om naar een feest, theater, bioscoop, concert, musea of bijeenkomst te gaan?; kunt u sporten?
0 = niet van toepassing
1 = ik kan geen één van deze aktiviteiten uitvoeren
2 = ik voer slechts een deel van deze aktiviteiten uit; meestal heb ik de hulp van iemand anders nodig
3 = ik voer de meeste van deze aktiviteiten (grotendeels) zelfstandig uit
4 = ik voer al deze aktiviteiten volledig zelfstandig uit

8. Auto rijden/de bus nemen/fietsen
Bent u in staat om auto te rijden, de bus/metro te nemen en/of te fietsen?
0 = niet van toepassing
1 = ik kan geen één van deze aktiviteiten uitvoeren
2 = ik kan slechts één van deze aktiviteiten uitvoeren (eventueel met de hulp van iemand anders)
3 = ik kan twee van deze aktiviteiten uitvoeren (eventueel met de hulp van iemand anders)
4 = ik voer deze aktiviteiten volledig zelfstandig uit

9. Werk/studie
Bent u in staat uw eerdere (vóór het ziek worden) baan/studie uit te voeren?
0 = niet van toepassing
1 = ik ben niet in staat tot het uitvoeren van mijn eerdere baan/studie
2 = ik ben in staat tot het (gedeeltelijk) uitvoeren van aangepast(e) werk/studie
3 = ik ben in staat tot een gedeeltelijk uitvoeren van mijn eerdere baan/studie
4 = ik voer mijn eerdere baan/studie volledig uit

Appendix IV

Quality of life measure

Medical Outcome Study Short-form 36 items health survey (SF-36)

The ***Dutch version*** was applied in the current study. *Published with permission from the Iqola group.*

SF- 36 GEZONDHEIDSTOESTAND VRAGENLIJST

INSTRUCTIES: Deze vragenlijst gaat over uw standpunten t.a.v. uw gezondheid. Met behulp van deze gegevens kan worden bijgehouden hoe u zich voelt en hoe goed u in staat bent uw gebruikelijke bezigheden uit te voeren.
Beantwoord elke vraag door het antwoord op de aangegeven wijze te markeren. Als u niet zeker weet hoe u een vraag moet beantwoorden, geef dan het best mogelijke antwoord.

1. Hoe zou u over het algemeen uw gezondheid noemen:

(omcirkel één cijfer)

Uitstekend 1

Zeer goed 2

Goed 3

Matig 4

Slecht 5

2. Hoe beoordeelt u nu uw gezondheid over het algemeen, vergeleken met een jaar geleden?

(omcirkel één cijfer)

Veel beter nu dan een jaar geleden 1

Wat beter nu dan een jaar geleden 2

Ongeveer hetzelfde nu als een jaar geleden 3

Wat slechter nu dan een jaar geleden 4

Veel slecther nu dan een jaar geleden 5

3. De volgende vragen gaan over bezigheden die u misschien doet op een doorsnee dag. Wordt u door uw gezondheid op dit moment beperkt bij deze bezigheden? Zo ja, in welke mate?

(omcirkel één cijfer op elke regel)

BEZIGHEDEN	Ja, Ernstig beperkt	Ja, een beetje beperkt	Nee, helemaal niet beperkt
a. **Forse inspanning**, zoals hardlopen, tillen van zware voorwerpen, een veeleisende sport beoefenen	1	2	3
b. **Matige inspanning**, zoals een tafel verplaatsen, stofzuigen, zwemmen of fietsen	1	2	3
c. Boodschappen tillen of dragen	1	2	3
d. **Een paar** trappen oplopen	1	2	3
e. **Één** trap oplopen	1	2	3
f. Bukken, knielen of hurken	1	2	3
g. **Meer dan een kilometer** lopen	1	2	3
h. **Een paar honderd meter** lopen	1	2	3
i. Ongeveer **honderd meter** lopen	1	2	3
j. Uzelf wassen of aankleden	1	2	3

4. Heeft u in de afgelopen 4 weken, een van de volgende problemen bij uw werk of andere dagelijkse bezigheden gehad, ten gevolge van uw lichamelijke gezondheid?

(omcirkel één cijfer op elke regel)

	YES	NO
a. U besteedde **minder tijd** aan werk of andere bezigheden	1	2
b. U heeft **minder bereikt** dan u zou willen	1	2
c. U was beperkt in het **soort** werk of andere bezigheden	1	2
d. U had **moeite** om uw werk of andere bezigheden uit te voeren (het kostte u bv. extra inspanning)	1	2

5. Heeft u in de afgelopen 4 weken, een van de volgende problemen ondervonden bij uw werk of andere dagelijkse bezigheden ten gevolge van emotionele problemen (zoals depressieve of angstige gevoelens)?

(omcirkel één cijfer op elke regel)

	JA	NEE
a. U besteedde **minder tijd** aan werk of andere bezigheden	1	2
b. U heeft **minder bereikt** dan u zou willen	1	2
c. U deed uw werk of andere bezigheden niet zo **zorgvuldig** als gewoonlijk	1	2

6. In hoeverre hebben uw lichamelijke gezondheid of emotionele problemen u gedurende de afgelopen 4 weken gehinderd in uw normale omgang met familie, vrienden of buren, of bij activiteiten in groepsverband?

(omcirkel één cijfer)

Helemaal niet 1

Enigszins 2

Nogal 3

Veel 4

Heel erg veel 5

7. Hoeveel lichamelijke pijn heeft u de afgelopen 4 weken gehad?

(omcirkel één cijfer)

Geen 1

Heel licht 2

Licht 3

Nogal 4

Ernstig 5

Heel ernstig 6

8. In welke mate bent u afgelopen 4 weken door pijn gehinderd in uw normale werk (zowel werk buitenshuis als huishoudelijk werk)?

(omcirkel één cijfer)

Helemaal niet 1

Een klein beetje 2

Nogal 3

Veel 4

Heel erg veel 5

9. Deze vragen gaan over hoe u zich voelt en hoe het met u ging in de afgelopen 4 weken. Wil a.u.b. bij elke vraag het antwoord geven dat het best benadert hoe u zich voelde. Hoe vaak gedurende de afgelopen 4 weken –

(omcirkel één cijfer op elke regel)

	Altijd	**Meestal**	**Vaak**	**Soms**	**Zelden**	**Nooit**
a. Voelde u zich levelslustig?	1	2	3	4	5	6
b. Was u erg zenuwachtig?	1	2	3	4	5	6
c. Zat u zo in de put dat niets u kon opvrolijken?	1	2	3	4	5	6
d. Veolde u zich rustig en tevreden?	1	2	3	4	5	6
e. Had u veel energie?	1	2	3	4	5	6
f. Voelde u zich somber en neerslachtig?	1	2	3	4	5	6
g. Voelde u zich uitgeput?	1	2	3	4	5	6
h. Was u een gelukkig mens?	1	2	3	4	5	6
i. Voelde u zich moe?	1	2	3	4	5	6

10. Hoe vaak hebben uw lichamelijke gezondheid of emotionele problemen u gedurende de afgelopen 4 weken gehinderd bij uw sociale activiteiten (zoals vrienden of familie bezoeken, etc.)?

(omcirkel één cijfer)

Atijd 1

Meestal 2

Soms 3

Zelden 4

Nooit 5

11. Hoe JUIST of ONJUIST is elk van de volgende uitspraken voor u?

(omcirkel één cijfer op elke regel)

	Volkomen juist	Grotendeels juist	Weet ik niet	Grotendeels onjuist	Volkomen onjuist
a. Ik lijk wat gemakkelijker ziek te worden dan andere mensen	1	2	3	4	5
b. Ik ben even gezond als andere mensen die ik ken	1	2	3	4	5
c. Ik verwacht dat mijn gezondheid achteruit zal gaan	1	2	3	4	5
d. Mijn gezondheid is uitstekend	1	2	3	4	5