Aus dem Institut für Gesundheits- und Rehabilitationswissenschaften
der Ludwig-Maximilians-Universität München
Vorstand: Prof. Dr. med. Gerold Stucki

Connecting perspectives on stroke disability:

The measurement and the classification approach

Dissertation
zum Erweb des Doktorgrades der Humanbiologie
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität zu München

vorgelegt von
Szilvia Geyh
aus
Budapest, Ungarn
2007

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter:	Prof. Dr. med. Gerold Stucki
Mitberichterstatter:	Prof. Dr. R. Haberl
	Priv. Doz. Dr. Dr. R. Werth
Mitbetreuung durch den promovierten Mitarbeiter:	Dr. rer. biol. hum. Alarcos Cieza
Dekan:	Prof. Dr. med. Dietrich Reinhardt
Tag der mündlichen Prüfung:	13.02.2007

I thank Professor Gerold Stucki and Dr. Alarcos Cieza for their support, council, and inspiration. Also, I would like to thank the team of the ICF Research Branch for a working atmosphere of mutual appreciation and friendship. I owe many thanks to my family for giving me the strength to cope with all difficulties along the way, but most of all to my husband Christian Geyh for being my vital source of energy.

CONTENTS

I. BACKGROUND

I.1. Stroke and disability

Stroke is a frequently occurring condition and a common cause of death and disability. Stroke was the 3rd leading cause of mortality in 2002 accounting for approximately 5.5 million deaths worldwide.[1] About 700,000 people experience a stroke each year in the United States.[2] In European and Asian populations, average annual stroke attack rates range from 185 to 638 per 100,000.[3] In Germany, the age and gender adjusted stroke incidence rate lies at 182 cases per 100,000 population per year.[4] However, secular trends in stroke mortality show a substantial decline in mortality rates from 79 to 29 deaths per 100,000 population between 1971 and 1994, while the number of stroke survivors increased from 1.5 to 2.4 million from 1973 to 1991 in the United States.[5] In Germany, more than 60%, across other populations 40% to 77% of patients survive beyond one year post-stroke.[6]

Many survivors are facing long-term ***disability***. Following stroke, each year 5 million people, corresponding to about one third of all incident stroke cases, are left permanently disabled according to the estimation of the World Health Organization. In the WHO Burden of Disease Study, cerebrovascular diseases were found to be the 3rd leading cause of lost 'disability-adjusted life years' (DALYs) in the developed countries, worldwide accounting for about 40 million DALYs in 1990.[7] Stroke is the biggest single cause of major disability in the United Kingdom. In the United States, about 1,160,000 non-institutionalized adults suffered from disability due to stroke in 1999.[8] Stroke and stroke related disability also imposes substantial economic burden to the patients, their families and the community, with direct and indirect costs of stroke being estimated at $51.2 billion for 2003 in the United States.

Stroke is defined as acute neurologic dysfunction of vascular origin with rapid onset of symptoms according to the affected regions of the brain.[9] Clinical manifestation of stroke can be described in terms of different arterial syndromes and varies depending on several factors including etiology, localization, and initial stroke severity, but also underlies considerable changes in the course of time.[10] Thus, acute symptoms often differ from the later picture of survivors' disability outcome.

The consequences of stroke on patients' functioning are usually complex and heterogeneous. Stroke has not only an impact on neurological functions (e.g. motor and sensory dysfunction), but may also leave survivors dependent in activities of daily living (ADL)[11,12] and leads to difficulties in patients' cognitive and mental state (e.g. attention, memory, language deficits, post-stroke depression, etc.).[13,14,15,16] In the Auckland Stroke Study, 61% of the patients with stroke reported 6 years after the acute event that they did not fully recover from stroke, and they were found to be at a substantially higher risk of being dependent in basic ADLs than age- and sex-matched controls.[12]

I.2. Two approaches to describe disability

Precise knowledge of patients' stroke related disability is necessary in health services provision and research. Clinical stroke management, but also epidemiological and clinical research, depend on the careful detection of functioning problems, as well as resources, in patients with stroke.

Two conceptual approaches to describe patients' burden, functioning and health can be distinguished: The health status measurement and the classification approach.[17] In the field of ***health status measurement***, the quantitative operationalization of patients' functioning takes the center stage, typically utilizing an abundance of available instruments with focus on specific aspects of functioning and based upon heterogeneous conceptual frameworks. In contrast, ***the classification approach*** is characterized by the comprehensive conceptualization and the qualitative representation of the full range of patients' functioning and health in terms of a systematic taxonomy.

I.2.1. The health status measurement approach

Health status measures, like standardized performance tests, rating scales, and questionnaires are used to assess patients' burden of disease, functioning and health. They yield comparable and easily communicable results in the form of profile or summary scores. Results from health status measurements allow for comparisons of an individual's state with population or reference group norms, as well as for comparisons across diverse populations, conditions, interventions, settings, or different time points. Health status measures can highlight target areas for necessary interventions, detect expected or unexpected changes, discriminate patient groups, can be useful to explain or predict health states, and may allow conclusions on the effectiveness, efficacy, safety or benefit of treatments.[18]

Health status measures are applied for a great variety of purposes in clinical, research, management, and policy settings. Following stroke, health status measures might be used for the examination and description of stroke impact, for monitoring, intervention evaluation, quality management, surveys, for individual as well as macro level health care planning and decision making.

Corresponding to the variety of application fields and measurement purposes a vast number of health status measures is available. Types of health status measures can be differentiated in various ways, for example, adopting the frequently used typology of generic, condition-specific and domain-specific measures.[19]

In the field of stroke, numerous measures exist to assess the wide scope of the event's impact and outcome. Several reviews provide an overview on these measures.[20,21,22,23,24] Bowling[20] describes various condition-specific measures, e.g., the National Institute of Health Stroke Scale (NIHSS),[25] as well as domain-specific instruments, e.g., the Mini-Mental State Examination (MMSE).[26] In more recent reviews[21,22,23] also generic health status measures used in stroke research are evaluated, for example the Medical Outcome Study Short-Form-36 Health Survey (SF-36),[27] the European Quality of Life Instrument (EuroQol),[28] and the COOP Charts.[29]

Most frequently, the effects of stroke are assessed by methods, like health professional ratings and performance tests,[30,31,32] e.g. the Barthel Index[33] or the MMSE. However, stroke survivors' everyday lives are affected in a variety of ways not easily captured by this type of methods.[34]

Quality of life (QoL) measures provide a comprehensive patient-centered approach to specify consequences of stroke. They are used to gather information not only on the disease, but also on the affected individual and his or her health experience.[35] Their use reflects the awareness that the patients' perspective is at the core of health care provision and research. Recently, several studies report on quality of life following stroke. Lower levels of QoL in stroke patients compared to healthy controls or general population norms have been found. [12,36,37] Further results also indicate a deterioration of QoL in the aftermaths of stroke.[38,39]

While in several diseases, like cancer, musculoskeletal, or cardiovascular conditions patient-centered measures are established,[40,41] the application of this type of measures in the field of stroke is still subject of current discussion for a variety of possible reasons.[31,32] The use of self or interviewer administered questionnaires is often difficult in stroke patients due to the deficits imposed by stroke itself, especially communication and cognitive impairments. Generic measures most frequently used to assess QoL in stroke[30,31,32] provide major advantages, especially for comparisons across different health conditions. However, they also have been criticized to have shortcomings that hamper their application with regard to the special population of stroke survivors. Generic measures may fail to cover patient problems that are of high importance in stroke, for example cognitive and language problems.[23,42] Additionally, as they might suffer from floor and ceiling effects and might be irresponsive to change, they cannot represent the diversity in stroke severity or the highly dynamic process of recovery.[21,23,43] Also, no generic measure exists, that has been validated for use in cognitively impaired patients.[44]

However, to meet these challenges recently new stroke-specific health status measures incorporating the patients' perspective are increasingly being developed.[45,46,47,48] They are expected to play an important role in future stroke measurement, as in management, epidemiological research, clinical, and drug trials QoL assessment is required to comprehensively capture stroke outcome.

I.2.2. The classification approach

I.2.2.1. International Classification of Functioning, Disability and Health (ICF) of the World Health Organization

The ***classification approach*** towards the description of patients' health state is represented by the World Health Organization's ***International Classification of Functioning, Disability and Health (ICF).***[49] The ICF provides a comprehensive conceptual framework and a unified standardized language to describe health and health related states, both at the individual, as well as at population levels. The ICF has been developed to complement the diagnostic information provided by the *International Statistical Classification of Diseases and Related Health Problems (ICD-10)*.[50]

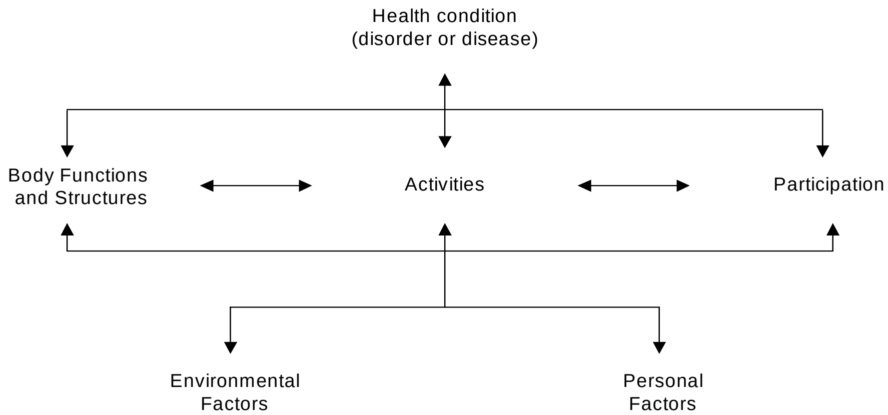

Figure 1. The biopsychosocial model of functioning, disability and health.

The ICF is based upon a ***biopsychosocial model***, which aims to integrate different perspectives of health into one unified and coherent view. The model relates to each other six components of health: the health condition, Body Functions and Structures, Activity, Participation, Environmental Factors, and Personal Factors.

The central concepts within this biopsychosocial model are functioning and disability. Functioning is an umbrella term for intact Body Functions and Structures, Activities and Participation. Functioning denotes the positive or neutral outcome of the bidirectional complex interaction between an individual with a health condition and his or her context. The complementary term disability is an umbrella term to denote impairments of Body Functions and Structures, Activity limitations and Participation restrictions. Disability is the negative outcome of the interaction between an individual with a health condition and his or her context. The current understanding of the interactions of the components of functioning, disability and health within a biopsychosocial approach is depicted in Figure 1.[49]

The ***ICF as a classification*** reflects the underlying biopsychosocial model. The ICF provides a list of Body Functions, a list of Body Structures, a joint list of Activities and Participations, and a list of Environmental Factors. Personal Factors are not implemented as a part of the classification, yet. Moreover, health conditions are not classified by the ICF, but are classified by an other member of the WHO's family of international classifications, the ICD-10. The ICF as a classification is a listing of categories, which are the units of the classification. The categories are organized within a hierarchically nested structure. An example from the component Body Functions is presented in the following:

b1	*Mental functions*	(first/ chapter level)
b114	*Orientation functions*	(second level)
b1142	*Orientation to person*	(third level)
b11420	*Orientation to self*	(fourth level)
b11421	*Orientation to others*	(fourth level).

The categories are accompanied by definitions, examples, inclusion, and exclusion criteria.

The endorsement of the ICF by the Word Health Assembly in May 2001 marks an important milestone in health services provision and research, and especially in the field of rehabilitation.[51] The ***potential uses*** of the ICF are numerous. The ICF provides a universal terminology to describe functioning and disability, which is applicable independent of a specific disease or health condition, etiology and pathogenesis, of the profession or specialization of the user, of time, place, culture, country, or health care system. The ICF is a multipurpose flexible tool that allows describing health in individuals or groups, comparing different health conditions, persons, defined groups, time points, countries, health care systems. It represents a useful common platform to communicate in a multi-professional team, between different departments or facilities, between clinicians and scientists, politicians, decision-makers, and not least, to communicate with the patients.[49,52] The ICF can be used in clinical practice and rehabilitation to structure and to lead through, thus to facilitate the rehabilitation process. It can provide a standardized frame for rehabilitation understood as a problem solving process with its steps: assessment and goal setting, assignment, intervention, and evaluation.[17,53,54] It can be used in teaching and education of health professionals,[55,56] but also to aggregate information, e.g. for health reporting purposes, public health information systems and epidemiology to build the necessary evidence basis for individual clinical, population-based institutional, or political decisions.[57] Also, the ICF is a useful tool for research, e.g. to select and to describe study populations, and also as a heuristic tool to clarify concepts, to generate and test hypotheses, or to explain health states.

Actual ***current uses*** of the ICF can only be exemplified as they are too numerous despite the recent introduction of the ICF. In the area of legislation and policy, Germany is among the first countries to have established the ICF. The ICF is an important pillar for the conceptualization of disability and the provision of rehabilitation according to the current social security codes, as well as according to mandatory guidelines based on them (e.g. SGB V, §92 Abs. 1 Satz 2 Nr 8; SGB IX, §4; Richtlinien des Gemeinsamen Bundesausschusses über Leistungen zur medizinischen Rehabilitation). But also in other countries, e.g. the United States, Canada, Australia, the Netherlands, France, and Japan, the ICF has been used in legislation or other ongoing governmental activities, while several additional countries plan to use and to implement the ICF, e.g. for national surveys, health statistics and information systems.[58,59] Efforts to apply the ICF in the clinical setting and in research are made in several countries, across different health professional groups, for different health conditions. The ICF has proved its usefulness especially in the area of ***neurorehabilitation***. It is a fruitful tool for structuring the process of neurorehabilitation, for goal setting and assessment, for documentation, reporting, and multidisciplinary team communication.[54,60,61,62,63]

I.2.2.2. The development of ICF Core Sets

The ICF is a highly comprehensive classification containing more than 1400 categories to describe patient's functioning, disability, and health. This comprehensiveness is a major advantage and strength of the ICF. But at the same time it is the major challenge to its practicability and feasibility.

To enhance the applicability of the classification, ICF-based tools need to be tailored to the needs of the users, without forging the strengths of the ICF.[64] One approach to enhance the application of the ICF is the development of ***ICF Core Sets*** for specific health conditions.[51,65] Within this approach functioning and disability are explicitly connected to a defined health condition. This accords with the biopsychosocial model and with the requirement of the joint use of the ICF together with the ICD, as intended by the WHO. The WHO has recognized that in everyday clinical practice, only a fraction out of the total number of the ICF's categories will be needed.[66]

Accordingly, ICF Core Sets are practical tools that represent a selection of categories out of the whole classification. ICF Core Sets for specific health conditions are short lists of such ICF categories that are relevant to most patients with the respective condition.[64,65]

Scientifically based internationally agreed ***ICF Core Sets for 12 chronic health conditions*** have been developed in a collaborative project of the Ludwig-Maximilian University, Munich with the Classification, Assessment, Surveys and Terminology Group (CAS) of the WHO, and together with partner organizations worldwide, for the following chronic conditions:

- Breast Cancer[67]
- Chronic Ischemic Heart Disease[68]
- Chronic Widespread Pain[69]
- Depression[70]
- Diabetes Mellitus[71]
- Low Back Pain[72]
- Obesity[73]
- Obstructive Pulmonary Diseases[74]
- Osteoarthritis[75]
- Osteoporosis[76]
- Rheumatoid Arthritis[77]
- Stroke[78]

For each of these conditions two types of ICF Core Sets have been developed. ***Comprehensive ICF Core Sets*** include the prototypical spectrum of problems in functioning in patients with a specific condition. They have been developed to guide multi-professional comprehensive assessment and to include as few as possible, but as many as necessary ICF categories to sufficiently describe patients' functioning. The ***Brief ICF Core Sets*** can serve as minimum data sets to be reported in every clinical study and to be assessed at any clinical encounter involving patients with the specific health condition. They include the most important categories in any situation, setting, country or culture.[64,65] Using the universal terminology of the ICF, ICF Core Sets preserve all advantages and potentials of the classification, at the same time by their manageable size enhancing its feasibility for the application field of a particular health condition.

The development of ICF Core Sets for 12 chronic health conditions is conceived as an evidence-based scientific process and at the same time as a consensus process. Preliminary studies have been conducted to provide the evidence basis for selecting the relevant categories for the ICF Core Sets. The preliminary studies for each health condition included a Delphi exercise to represent the health professionals' perspective,[79] a systematic review on outcomes used in randomized clinical trials (RCTs) to represent the researchers' perspective,[30,80,81,82,83] and an empirical data collection based on the ICF Checklist[84] representing the perspective of patients undergoing inpatient or outpatient rehabilitation.[85] The ICF categories to be included in the first versions of the ICF Core Sets were identified in international consensus conferences by the means of a formal decision-making and consensus process integrating the evidence gathered from the preliminary studies.

Currently, the ***validation*** of the first ICF Core Sets for chronic health conditions is underway. In the studies of this validation phase, alike the preliminary studies, quantitative and qualitative methodological approaches complement each other. The first ICF Core Sets are validated from the patients' perspective, the health professionals' perspective, and using empiric data collected in an international multicentric study. The results of the ICF Core Sets validation studies will be presented by the end of 2007 at a WHO ICF conference. Representatives from the different WHO world regions will then decide on the final endorsement of the ICF Core Sets for chronic health conditions.

I.2.2.3. ICF Core Sets for Stroke

The *ICF Core Sets for Stroke* are selections of salient ICF categories out of the whole classification, which describe the spectrum of problems in stroke patients' functioning based on the universal language of the ICF. The *ICF Core Sets for Stroke* represent the practical implementation of the classification approach in clinical practice and research to describe stroke related disability.

The total number of categories in the ***Comprehensive ICF Core Set for Stroke*** is 130. No categories at the 3rd and 4th levels of the classification are included, but all categories belong to the 2nd level of the ICF. The Comprehensive ICF Core Set contains 41 categories (32%) from the component Body Functions, 5 categories (4%) from the component Body Structures, 51 (39%) from the component Activities and Participation, and 33 (25%) from the component Environmental Factors. The *Comprehensive ICF Core Set for Stroke* is the largest of the ICF Core Sets developed for the 12 most burdensome chronic conditions. The large scope of categories included in the Comprehensive ICF Core Set reflects the important and complex impairments, limitations, and restrictions of Activities and Participation involved, as well as the numerous interactions with Environmental Factors.

The ***Brief ICF Core Set for Stroke*** comprises a total of 18 categories, which represent 14% of the categories chosen in the Comprehensive ICF Core Set. The Brief ICF Core Set includes 6 categories from the component Body Functions (15%), 2 categories (40%) from Body Structures, 7 (14%) from Activities and Participation, and 3 (9%) from Environmental Factors. The selected categories for the Brief ICF Core Set account for the fundamental and most striking aspects of stroke related functioning.

The ICF categories included in the *Comprehensive ICF Core Set for Stroke* are shown in Appendix 1. Appendix 2 shows the ICF categories that have been selected for the *Brief ICF Core Set for Stroke*.

I.3. The linking method

I.3.1. The linking method as connecting approach

The two perspectives on patients' burden, functioning and health, ***the health status measurement and the classification approach can be regarded as complementary principles.*** While from the classification perspective the ICF and the ICF Core Sets can serve as standards to define ***what*** to measure, from the perspective of health status measurement the question ***how*** to measure can be answered.

In the field of stroke, first steps have already been made to connect health status measurement and the ICF. Currently, several instruments have been developed that are based on the ICF's biopsychosocial framework, or on its predecessor, the *International Classification of Impairments, Disabilities and Handicaps* (ICIDH)[86] (e.g. the Burden of Stroke Scale,[42,47] or the Stroke Impact Scale[87]). The ICF's biopsychosocial model has also been regarded useful in measurement reviews to organize a variety of different instruments according to their content and thus to facilitate the choice of instruments in stroke rehabilitation.[24,88,89,90]

An explicit connection between the two perspectives is established by the so-called ***linking method***.[91,92] In the same way like the ICF can be used to describe patients' functioning and health, it can also be used to describe and to compare health status measures. Using the ICF's well defined categories as an a priori, independent, and external reference system to represent the contents of measures, it is possible to explore and compare these contents in a comprehensive, standardized, transparent and straightforward way. The linking method relies on the smallest possible units of content, namely on single concepts within the items. Therefore, it gives a fine-grained, clear, and precise picture of the addressed contents of the instruments.

The linking method includes the identification of single content concepts within the items of the candidate measures, and the linking of the content concepts to the corresponding ICF categories, which most precisely represent the concept. The linking is performed utilizing a set of established linking rules.[91,92] The linking rules are guidelines, which enable concepts contained in health status measures to be translated into the language of the ICF in a standardized manner. Based on the linking method it is possible to specify, quantify, and compare the contents of different health status measures.

I.3.2. The linking method as basis to study the content validity of health status measures

Corresponding to the variety of application fields and measurement purposes a vast number of health status measures exists,[93] and new versions of old instruments are continuously proliferating. The large number of available instruments poses a growing challenge to clinicians and researchers when it is to select the appropriate measure for a given application situation. The selection of measures is an essential step to be taken in planning any data collection. A thorough examination and evaluation is necessary to fulfill the aim of sound instrument selection.[94,95] The psychometric properties (e.g. reliability, validity, etc.) and the application related features (e.g. administration mode, burden, etc.) of the instruments need to be accounted for. However, the first and most important concern is ***content validity***: "Does the measure cover the contents to be measured?"

A new way of examining health status measures' content validity arises from the application of the linking method along with the ICF Core Sets. The *ICF Core Sets for Stroke* can be used as common reference to define "what to measure" in stroke. Applying the linking method, the contents of health status measures can be translated into lists of ICF categories. In this way, a direct and straightforward comparison of the measures with the ICF Core Sets is possible. The extent to which a health status measure's contents cover the spectrum of the ICF Core Sets' categories can serve as an indicator for content validity.

I.4. Rasch analyses as a method to study the psychometric properties of health status measures

Beyond content validity, meaningful measurement essentially depends on the psychometric quality of the applied instruments. Techniques based on modern test theory, especially ***Rasch analyses***, are increasingly adopted to ensure instruments' psychometric properties. Rasch analysis, first applied in educational settings[96] gained in importance within the field of health status measurement in the past 20 years and entails enthusiasm and great expectations.[97,98,99] Rasch analysis is currently the only way to achieve objective measurement in the human sciences.[100]

Objective measurement implies the measurement of one single attribute at a time, the units of measurement maintaining equal size across the measurement continuum, the result of measurement being independent of the particular instrument used, of the person measured, as well as of the person measuring.[100,101] Only objective measurement yields interval scale data, which are additive, reproducible, comparable, and suitable for further analyses, which cannot be applied to ordinal scale raw scores. In summary, diagnosis and evaluation can be fundamentally improved by the merits of objective measurement.

The Rasch method has already been applied for different purposes in the field of stroke health status measurement in the past 10 years. The Rasch method has been used in the development of new instruments.[102,103,104] Also, it has been applied to ensure the measurement properties of existing instruments and to transform their raw scores into linear measures, e.g. for the Functional Independence Measure (FIM),[105,106] the Barthel Index (BI),[33,107] and the National Institutes of Health Stroke Scale (NIHSS).[25,108] Rasch analysis techniques proved to be especially useful in addressing the issues of the cross-cultural validity and comparability of instruments[109,110] and in the psychometric evaluation of measures in different countries following their translation.[111,112,113,114] In the future, Rasch based methods are expected to gain in importance in stroke health status measurement.

II. RESEARCH OBJECTIVES

In the following, the doctoral thesis is subdivided into four parts. The first three parts present different studies performed to pursue the objectives named below. Each of the three studies contains a respective discussion section referring to the results of the study. The fourth part of the doctoral thesis refers to aim four, namely the discussion of the relationship between the methods presented in the previous three parts. Thus, the current doctoral thesis aims:

1. To illustrate, how the connection between the health status measurement approach and the classification approach can be established. For this purpose the linking method is applied to examine and to compare the contents of patient-centered health status measures used in stroke based on the ICF.

2. To demonstrate, how this approach can be used to select health status measures based on their content validity. For this purpose, a comparison of stroke-specific health status measures with the *Comprehensive ICF Core Set for Stroke* is conducted.

3. To show, how the psychometric features of health status measures can be examined based on modern test theory and Rasch based methods. For this purpose the psychometric evaluation of a selected measure is performed using Rasch analyses.

4. To discuss the relationship between the demonstrated methods in the context of the connection of the health status measurement and the classification approach.

III. APPLYING THE LINKING METHOD:

Content comparison of patient-centered health status measures used in stroke based on the International Classification of Functioning, Disability and Health (ICF)

III.1. Specific aims

The general aim of this study is to illustrate how the connection between the health status measurement approach and the classification approach can be established. The specific aims are: (1) to identify current generic and condition-specific patient-centered health status measures applied in stroke patients, (2) to examine the contents of the single measures based on their linking to the ICF, and (3) to examine and compare the contents of generic and stroke-specific measures.

III.2. Methods

A systematic literature review was conducted to identify and select current generic and condition-specific patient-centered health status measures applied in stroke. The contents of the selected measures were examined by extracting the single concepts from the measures' items and by linking them to the ICF using established linking rules.[91,92] The frequencies of ICF categories representing the instruments' concepts built the basis of the quantitative descriptive analysis and content comparison.

III.2.1. Systematic literature review

The electronic databases MEDLINE, EMBASE and PsycINFO were searched using the keywords ‘cerebrovascular accident’ or ‘stroke’, and ‘health status’ or ‘quality of life’. The exact search terms vary by database, as the specific thesaurus vocabulary of the given database was used. Searches were limited to original articles published between 1999 and 2004 in English language. No restrictions regarding the study design were imposed at this point.

Study eligibility was checked for in three steps. Using the information displayed by the main abstract, in a first step, descriptive, evaluative or psychometric studies generating firsthand data about patients with ischemic or hemorrhagic stroke were included. Reviews, case reports, economic evaluations or primary prevention studies were excluded, as well as studies including non-stroke or healthy persons. In a second step, studies were selected that report the use of multidimensional patient-centered health status measures. Finally, full text articles of the selected studies were retrieved and checked again using the same eligibility criteria. All patient-centered health status measures used in the included studies were documented. The five most frequent generic and the five most frequent stroke-specific instruments were to be selected according to their rank order for content examination and comparison.

III.2.2. Linking of measures to the ICF

The selected measures were linked to the ICF by two trained health professionals utilizing a modified version of established linking rules.[91,92] Both linkers conducted this procedure independently from each other, thus two independent linking versions of each instrument were created, then compared. Linker disagreement was resolved by structured discussion and informed decision of a third expert.

The linking rules are guidelines, which enable concepts contained in health status measures to be linked to the ICF in a standardized manner.[91,92] The linking procedure includes the identification of single content concepts within the items of the measures, and the linking of these content concepts to the corresponding ICF categories, which most precisely represent the concepts. If an item of a measure contains more than one concept, each concept is linked separately. For example, the item of the Burden of Stroke Scale[42,47] “Because of your stroke, how difficult is it for you to kneel down or bend over?” has been linked to the ICF categories *d4102 Kneeling* and *d4105 Bending*.

Concepts that cannot be linked to the ICF are documented in two ways. If a concept is not sufficiently specified to make a decision about which ICF category to use, the concept is coded ‘nd’ or ‘not definable’. For example, concepts such as “physical disability”, or “health” are not sufficiently specified for precise linking. If a concept is not represented by the ICF, this concept is labeled ‘nc’ or ‘not covered’. Such concepts may be related for example to Personal Factors for which no categories currently exist, although they are considered a part of the contextual factors within the ICF’s biopsychosocial model. Also, ‘nc’ may represent concepts that lay outside the scope of the ICF, e.g. specifications of disease conditions or diagnoses.

III.2.3. Data analyses

The documentation of the linking process, data management and control were conducted using a simple database.[115] The reliability of the linking process was evaluated by calculating kappa coefficients[116] and nonparametric bootstrapped confidence intervals[117] based on the two independent linking versions of each instrument. Kappa statistics were calculated per component at the 1st, 2nd, and 3rd ICF levels to indicate the degree of agreement between the two health professionals conducting the linking procedure. The kappa analysis was performed with SAS.[118]

The quantitative descriptive analysis and content comparison of the measures was based on the final agreed version of the linking of each instrument. The *identified concepts* from the selected instruments are described according to their frequency distribution across ICF components and ICF levels of hierarchy. The number of concepts, which are not linked to the ICF and are denoted ‘not covered’ or ‘not definable’ is reported. For each instrument, the ratio of the number of concepts and the number of the instruments’ items is used as an index to characterize *content density*. Hereby, a value of 1 means, that each item of the instrument contains one concept. The higher the index value the more concepts are contained within one single item on average.

The number of *different ICF categories* applied is reported. Their frequency distribution across the ICF components, and their percentage based on the number of total existing ICF categories indicates the *bandwidth of content coverage* by the instruments. The different ICF categories’ frequency distribution across the ICF levels of hierarchy indicates the *specificity of the content* within the instruments. For each instrument, the ratio of the number of different ICF categories employed and the number of concepts is used as an index to characterize the *content diversity* of an instrument. A value of 1 indicates that each concept of the instrument is linked to a different ICF category. A value towards zero indicates that several concepts are linked to one and the same ICF category.

Each ICF category's frequency across the selected instruments is examined. ICF categories are counted once per item to obtain the frequency of an ICF category. The resulting value indicates the number of items within the intrument containing any concept represented by this ICF category. Frequencies are also summarized at the chapter level.

All the data reported are summarized for the generic and the stroke-specific measures, respectively.

III.3. Results

The electronic literature searches in MEDLINE, EMBASE and PsycINFO were conducted in May 2004 and yielded 1727 hits. After the three steps of the eligibility checks 71 studies have been included. Twenty-three different patient-centered multidimensional health status measures used in these studies were identified. Six generic and seven stroke-specific instruments were selected according to their frequency rank order for the ICF based content examination. Table I shows all identified measures and their frequency.

The six generic measures selected for ICF based content examination are the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36), the Reintegration to Normal Living Index[119] (RNL), the Sickness Impact Profile[120] (SIP), the European Quality of Life Instrument (EQ-5D), the London Handicap Scale[121] (LHS), and the Nottingham Health Profile[122] (NHP). The seven selected stroke-specific measures are the Stroke Impact Scale[48,87] (SIS), the Stroke-specific Quality of Life Scale (SSQoL), the Stroke and Aphasia Quality of Life Scale[123] (SAQoL-39), the Quality of Life Index - Stroke Version[124] (QLI-SV), the Stroke-Adapted Sickness Impact Profile-30[125] (SA-SIP-30), the Burden of Stroke Scale[42,47] (BOSS), and the Quality of Life Instrument for Young Hemorrhagic Stroke Patients[126] (HSQuale). Table II provides an overview on the major characteristics of the selected health status measures.

Table III shows the results of the evaluation of the linking procedure by kappa statistics and bootstrapped confidence intervals. Estimated kappa values range from 0.46 to 0.84. None of the 95% confidence intervals encloses zero, thus linker agreement exceeds chance.

Table IV shows the number of *identified concepts* from the selected instruments, as well as summarized for the generic and stroke-specific measures. The table also includes the concepts' distribution across the ICF components and the ICF levels of hierarchy, as well as the number of concepts not linked to the ICF. Within the 13 selected instruments 979 concepts have been identified, 441 within the generic and 538 within the stroke-specific measures. The SIP contains the highest (n=213) and the EQ-5D the lowest number of concepts (n=14). The measures' content density ratio shows the highest value for the LHS (9.2), which has 6 items containing 55 concepts. It is lowest for the QLI-SV (1.1), which has 72 items containing 76 concepts.

A total of 866 concepts were linked to the ICF. Most concepts have been linked to ICF categories from the component Activity and Participation (n=586, 60%). In contrast, 5% (n=47) of the concepts has been linked to Environmental Factors. Seven of the 13 instruments do not contain any concepts referring to this ICF component. No instrument contained concepts referring to Body Structures. Most concepts have been linked to ICF categories at the 2nd and 3rd levels of the ICF hierarchy (n=429, 44% and n=385, 39%).

113 concepts were not linked to the ICF. 58 concepts (6%) were 'not covered' by the ICF, 55 concepts were coded 'not definable' (6%). The QLI-SV contains the highest number of concepts that were classified as 'not covered' (n=20, 26%). The QLI-SV includes several items on overall satisfaction, for example, the item "How satisfied are you with your life in general?". Within these items the concept "satisfaction" is not covered by the ICF but represents a Personal Factor. The HSQuale contains the highest number of concepts (n=14, 11%) that were coded 'not definable'. This instrument includes for each of its subscales one overall question about the changes of quality of life, for example, the item "Would you say changes you have noticed in your physical functioning that have resulted from your bleed have increased, decreased, or not changed the quality of your life?". Within these items the concept "quality of life" cannot be mapped to any single specific ICF category and is coded 'not definable'.

Table V shows the number of *different ICF categories* employed to represent the instruments' concepts and their frequency distribution across the ICF components and the ICF levels of hierarchy for the selected instruments, as well as summarized for the generic and stroke-specific measures. A total of 200 different ICF categories corresponding to 14% of all existing ICF categories have been used to map the contents of the 13 instruments. Within the generic and stroke-specific measures 150 different ICF categories have been used for linking (10%), respectively. The bandwidth of content coverage is largest for the component Activity and Participation. From this ICF component 126 different categories were used, which cover 32% of all existing ICF categories of this component. The bandwidth of content coverage is smallest for the component Environment; 19 different categories have been applied covering 7% of all existing categories of the Environmental Factors. The instrument with the broadest bandwidth of content coverage is the SIP. To link the concepts of the SIP 104 different ICF categories have been used covering 5% of all Body Functions, 17% of all Activities and Participations, and 4% of the Environmental Factors. In contrast, the instrument with the narrowest bandwidth of content coverage is the EQ-5D. For the linking of the concepts of the EQ-5D 12 different ICF categories were necessary covering 3% of all Activities and Participations. The bandwidth of content coverage indicated by the percentage of ICF categories used across the ICF components is similar for generic and stroke-specific measures.

The content diversity ratio is lowest for the QLI-SV (0.33), where 25 different ICF categories were used to represent 76 concepts. It is highest in the EQ-5D (0.86), where 12 different ICF categories have been used to map the 14 concepts.

The different ICF categories applied to map the instruments' contents most frequently belong to the 3rd level (n=109) and the 2nd level (n=78) of the ICF hierarchy. The level of specificity indicated by the distribution of different ICF categories across the levels of hierarchy is similar for generic and for stroke-specific measures.

Tables VI to IX show the frequency of *each single ICF category* from the components Body Functions, Activity and Participation, and Environment used for the final linking of concepts from the selected measures, and summarized for the generic and stroke-specific instruments. No single category is contained in all instruments. The ICF categories *d540 Dressing* and *d760 Family relationships* are both represented in 11 of the 13 instruments, in all generic measures and in 5 of the stroke-specific ones. Dressing is not contained in the SIS and the QLI-SV. However, the SIS contains related categories, which are more specified. From the total of 200 different categories used 77 (40%) applied to only one of the 13 selected measures. Furthermore, 50 ICF categories (25%) are represented in the stroke-specific instruments only and 51 categories (25%) are addressed in generic measures only. For example, specified memory functions (*b11440, b1441, b1442*) or specific mental functions of language (*b167, b16700, b16701, b16710*) are only included in stroke-specific instruments, mainly in the SIS, the BOSS, and the HSQuale. On the other hand, 13 out of the total 19 different ICF categories from the component Environmental Factors are only included in generic measures, mainly in the RNL and the SIP.

Overall, the most frequently used category is *b152 Emotional functions*, which is contained in 53 items from 10 different instruments and has the highest frequency within the generic as well as within the stroke-specific measures. Beyond the most frequent category *b152 Emotional functions*, generic and stroke-specific measures differ with respect to the categories most frequently addressed. Within the generic measures the categories *d2102 Undertaking a single task independently* (n=15), *b280 Sensation of pain* (n=13), *d760 Family relationships* (n=13) were used most frequently. Within the stroke-specific measures the categories *d330 Speaking* (n=14), *d450 Walking* (n=11), and *b1300 Energy level* (n=11) were most frequently applied.

Tables VI to VIII also show the ICF categories' frequency pooled at the chapter level for the selected measures and also summarized for generic and stroke-specific instruments. Referring to the ICF component Body Functions (Table VI), categories from the chapter *b1 Mental functions* have been used most frequently to address the instruments' contents. *Mental functions* are contained in 159 items of the 13 selected measures, in 57 items from generic, and 102 items from stroke-specific instruments. Within the ICF component Activity and Participation (Table VII), categories from chapter *d4 Mobility* have been applied most frequently. *Mobility* is addressed in a total of 149 items of the selected measures, in 79 items from generic, and 70 items from stroke-specific measures. Concerning the ICF component Environmental Factors (Table VIII), categories from the chapter *e1 Products and technology* have been used most frequently to represent the instruments' concepts. Categories belonging to *Products and technology* are contained in 30 items of the 13 instruments, in 25 items from generic and 5 items from stroke-specific measures.

III.4. Discussion

The present study provides an overview and comparison of current patient-centered health status measures in stroke with respect to their covered contents using the ICF as independent, external reference system. The examination of the instruments' contents relies on the smallest possible units of content, namely on single concepts within the items, which gives a clear and precise picture of the addressed contents of the instruments and allows for straightforward comparisons. The results of the content comparison provide valuable information to facilitate and to account for the selection of appropriate instruments for specific purposes of data collection in clinical as well as research settings.

Although several reviews can be relied on that aim to facilitate the selection of appropriate instruments applied in stroke, they mainly describe the instruments psychometric properties.[20,22,23,24,43] Content comparisons are scarcely represented in the literature. Golomb et al. (2001) have conducted a comparison of health status measures' content coverage across 11 domains, which were supposed to be affected by stroke. The domains were identified by the authors by means of a literature review and were used as a basis for rating content coverage. This comparison provides a helpful overview on the content of health status measures used in stroke and gives preliminary directions for instruments' choice. However, the domains used as reference system for this comparison were not defined explicitly and they were derived for the limited purpose of instrument comparison within this single review. Moreover the domains' scope might be not as comprehensive as to allow the representation of the entire content of the measures as they cover only selected stroke-related domains. Despite the useful information provided by this review, a further in-depth examination of the instruments' content is still necessary to decide on the selection of a particular instrument. In contrast, the content comparison presented here is based on the ICF, which serves as a universally accepted, well defined, and standardized reference system that allows for a fine-grained exploration and comparison of all contents of the measures.

The examination of the 13 instruments' content structure revealed insights into the measures content density, content diversity, bandwidth of content coverage, and specificity of content, which are useful features for instrument selection. While the instruments differ by length and number of concepts contained, based on the index of content density the level of content complexity of the instruments can be compared. For example, for application in stroke survivors with cognitive and communication impairments, instruments with lower levels of complexity might be preferred (e.g. QLI-SV, SIS, SAQoL-39, BOSS).

While the instruments differ by the number of concepts they contain and the number of ICF categories used to map these concepts, the index of content diversity and the bandwidth of content coverage indicate in a comparable way the extent to which instruments are differentiated, broad or focused with respect to special health domains. Instruments with a lower index of content diversity (e.g. QLI-SV, BOSS, SF-36) might be more differentiated and fine grained, including several concepts related to the same topic, while measures with a high content diversity (e.g. EQ-5D, LHS, SAQoL-39) may address their topics in a less differentiated and more parsimonious way. Instruments with a smaller bandwidth of content coverage (e.g. EQ-5D, SF-36, QLI-SV) may be focused on few important contents, while measures with a greater bandwidth of content coverage contain items across a higher number of different health domains (e.g. SIP, HSQuale, SIS). Depending on the special purpose of the instruments' intended use a different type of instrument would be appropriate, e.g. for surveys or individual decision-making.

Examining the representation of the instruments' contents, it is remarkable that a high percentage of the used ICF categories (40%) applied to only one of the selected measures, i.e. there is a high number of topics, which are addressed in only one instrument and no other instrument includes them. In this way, some measures are uniquely appropriate for special purposes. For example, the only instrument addressing specific memory functions is the SIS, consequently, for the purpose of evaluating a memory training program using a patient-centered health status measure, the SIS might be a preferred choice. In a similar way, even the category with the overall highest frequency, namely *b152 Emotional functions* is not included in all instruments, for example, it is not addressed in the RNL and the LHS, which thus might fall out of the selection pool for purposes when emotions are of special interest. A further important finding of this study refers to the representation of Environmental Factors within the selected measures. Only few measures involve the influence of Environmental Factors, like assistive devices or support, e.g. the RNL or the SIP. However, for example in community rehabilitation settings or for use in the context of special health professions, e.g. in occupational therapy, instruments including Environmental Factors might be preferred.[127]

No systematic differences between generic and stroke-specific measures were found with regard to their content structure; however, there seem to be differences in the content representation. Stroke-specific measures more often address different mental functions, than generic measures. Thus, for purposes when mental functions following stroke are of special interest an appropriate instrument might rather be chosen from the pool of stroke-specific measures. In contrast, Environmental Factors are more often addressed in the selected generic instruments. However, referring to Activities and Participation no systematic difference between generic and stroke-specific measures is apparent.

Further differences between the contents of generic and the stroke-specific instruments have been found at the most frequently used ICF categories. Within the generic measures, independence, pain and family relations are addressed most often, which are the areas where the most burdensome patient problems may arise in any health condition. The most often addressed areas within the stroke-specific measures, i.e. walking, speaking, and energy, represent the direct impact of stroke on affected patients' lives. This finding provides support to the usefulness of the applied linking procedure as it clearly reflects the conceptual differences between the generic and stroke-specific instruments.

The current study is subject to several limitations. The systematic literature review used to identify current patient-centered health status measures in stroke relied on a simplified review methodology, using specific rather than sensitive search strategies and relying to a large extent on information contained in the paper abstracts. Still, compared with other reviews on measures in stroke[21,22,23,43] the instruments identified cover the most frequently used established patient-centered health status measures and in addition also include several most recently developed instruments.

The linking process has been evaluated by calculating kappa coefficients, which showed satisfactory results for linker agreement. Kappa is an often used and simple indicator of agreement accounting for chance. However, unsystematic error due to chance appears to be of secondary relevance for the linking procedure and further analyses, e.g. using modeling methods, would be useful in future to explain the disagreement between the linkers (e.g. due to experience or profession) to refine the linking method.

Finally, when interpreting the indices for content diversity, bandwidth of content coverage and specificity of content, it is implied that the ICF is the accepted reference with its given categories and its given levels of hierarchy as the units of comparison. Thus, the results of the content comparison of the selected instruments only hold relative to this frame of reference. However, the ICF is expected to become the one generally accepted framework to describe functioning and health in all health related fields.[128]

Table I: Frequency of the use of the 23 different multidimensional patient-centred health-status measures in 71 included studies listed for different study types

Instruments		**Rank order**	**Frequency in all study types n=71 (%)***	**Descriptive studies n=42**	**Intervention evaluation n=11**	**Psychometric studies n=18**
	Generic instruments					
SF-36	Medical Outcomes Study 36-Item Short-Form Health Survey	1	31 (44)	20	3	8
RNL	Reintegration to Normal Living Index	2	9 (13)	6	2	1
SIP	Sickness Impact Profile	2	9 (13)	7	1	1
EQ-5D	European Quality of Life Instrument	3	7 (10)	4	1	2
LHS	London Handicap Scale	4	5 (7)	2	1	2
NHP	Nottingham Health Profile	5	4 (6)	1	3	
COOP	Dartmouth COOP Charts	6	2 (3)		1	1
15-D	15-Dimensional Measure of Health Related Quality of Life Test	7	1 (1)		1	1
LIFE-H	Assessment of Life Habits	7	1 (1)	1		
AQoL	Assessment of Quality of Life	7	1 (1)			1
CHART	Craig Handicap Assessment and Reporting Technique	7	1 (1)			1
HUI II	Health Utilities Index Mark II	7	1 (1)	1		
HSQ	Health Status Questionnaire	7	1 (1)	1		
LQLP	Lancashire Quality of Life Profile	7	1 (1)	1		
QLI	Quality of Life Index	7	1 (1)	1		
WHOQOL	World Health Organization Quality of Life Scale	7	1 (1)			1
	Stroke-specific instruments					
SIS	Stroke Impact Scale	1	5 (7)	1		4
SSQOL	Stroke-Specific Quality of Life Scale	2	4 (6)	2		2
SAQOL-39	Stroke and Aphasia Quality of Life Scale	3	2 (3)	1		1
QLI-SV	Quality of Life Index - Stroke Version	3	2 (3)	2		
SA-SIP30	Stroke-Adapted Sickness Impact Profile-30	3	2 (3)	1	1	
BOSS	Burden of Stroke Scale	4	1 (1)			1
HSQuale	Quality of Life Instrument for Young Hemorrhagic Stroke Patients	5	1 (1)			1

* Percentages do not sum up to 100, as in several studies more than one instrument was applied.

Table II: Overview on the major characteristics of the selected generic and stroke specific patient-centred health status measures

	Instrument	Dimensions/ subscales	Mode of administration	Number of items and response options	Time frame	Time to complete	Reliability and validity in stroke examined	Versions	References
SF-36	**Medical Outcome Study Short-Form-36 Health Survey**	Physical functioning, mental health, social functioning, role limitations physical, role limitations emotional, general health perceptions, vitality, bodily pain	Self Interviewer Proxy Telephone	36 items yes/ no 3 to 6 point scales	Present Past 4 weeks	~10 min	yes		27, 129, 130, 131, 132
RNL	**Reintegration to Normal Living Index**	Daily functioning, perception of self	Self Interviewer Proxy	11 items VAS	Present	~10 min	yes		119, 133,134
SIP	**Sickness Impact Profile**	Sleep/ rest, eating, work, home management, recreation/ pastimes, ambulation, mobility, body care and movement, social interaction, alertness behaviour, emotional behaviour, communication	Self Interviewer	136 items yes/ no	Today	~20-30 min	yes	A stroke specific version exists: the SA-SIP30	120, 125, 135, 136
EQ-5D	**European Quality of Life Instrument**	Mobility, self-care, usual activities, pain/ discomfort, anxiety/ depression	Self Interviewer Proxy	5 items 3 ordered response options 100 point VAS	Today compared with past 12 months	~5 min	yes		28, 129, 137, 138, 139
LHS	**London Handicap Scale**	Mobility, physical independence, occupation, social integration, orientation, economic self-sufficiency	Self Interviewer	6 items 6 ordered response options	Last week	not reported	yes		121, 140, 141
NHP	**Nottingham Health Profile**	Physical abilities, pain, emotional reactions, energy level, sleep, social isolation	Self Interviewer	45 items yes/ no	Present	~10-15 min	yes		122, 142, 143, 144
SIS	**Stroke Impact Scale**	Strength, memory/ thinking, emotion, communication, ADL/ IADL, mobility, hand function, participation	Self Interviewer Proxy	59 items 5 point scale	Past 1 to 4 weeks	not reported	yes		48, 87, 145, 146, 147, 148
SSQol	**Stroke-Specific Quality of Life Scale**	Self-care, vision, language, mobility, work/ productivity, upper extremity function, thinking, personality, mood, family roles, social roles, energy	Interviewer Proxy	62 items 5 point scale	Past week	~12 min	yes	A modified version for aphasic patients exists: the SAQoL-39	45
SAQol -39	**Stroke and Aphasia Quality of Life Scale-39**	Physical, psychosocial, communication, energy	Interviewer	39 items 5 point scale	Past week	not reported	yes	Modified version of the SS-QoL for aphasic patients	123, 149
QLI-SV	**Quality of life Index Stroke Version**	Health and functioning, socio-economic, psychological-spiritual, family	Self Interviewer	72 items 6 point scale	Present	~10 min	yes	Stroke specific version of the QLI	124, 150, 151
SA-SIP30	**Stroke- Adapted Sickness Impact Profile-30**	Body care and movement, social interaction, mobility, communication, emotional behaviour, household management, alertness behaviour, ambulation	Self Interviewer	30 items yes/ no	Present	not reported	yes	Stroke specific version of the SIP	125, 152
BOSS	**Boss burden of stroke scale**	Mobility, self care, swallowing, energy and sleep, domain mood, domain satisfaction, domain restriction, positive mood, negative mood, communication, cognition, social relations	Interviewer	65 items 5 point scale	Present Since stroke	not reported	yes		42, 47,153
HS Quale	**Instrument for Young Hemorrhagic Stroke Patients**	General outlook, physical functioning, cognitive functioning, relationships, social/ leisure activities, emotional well-being, work/ financial status	Interviewer	54 items 4 to 7 ordered response options 3 open questions	Implicit present Since bleed	~15 min	yes		126

Table III: Kappa coefficients and nonparametric bootstrapped 95% confidence intervals at the component, chapter, 2nd and 3rd levels of the ICF for the three ICF components Body Function, Activity and Participation, and Environmental Factors

ICF Levels	ICF Components	Body Function	Activity and Participation	Environment
Components	0.81 [0.78;0.84]			
Chapter Level		0.67 [0.58;0.76]	0.84 [0.80;0.87]	0.62 [0.48;0.81]
2nd Level		0.76 [0.69;0.85]	0.78 [0.78;0.81]	0.46 [0.24;0.72]
3rd Level		0.67 [0.63;0.73]	0.78 [0.78;0.79]	0.51 [0.26;0.79]

Table IV: The number of identified concepts from the selected instruments, summarised for the generic and stroke specific measures. The table also includes the number of concepts as represented by ICF categories across the four ICF components and the ICF levels of hierarchy, as well as the number of concepts, which were not linked to the ICF.

	Total	**Generic**	**Specific**	**SF-36**	**RNL**	**SIP**	**EQ-5D**	**LHS**	**NHP**	**SIS**	**SSQoL**	**SAQoL -39**	**QLI-SV**	**SA-SIP30**	**BOSS**	**HS Quale**
	n (%)*	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***	**n (%)***
Number of items	620	239	381	36	11	136	5	6	45	59	62	39	72	30	65	54
Number of concepts																
Total	979	441	538	54	43	213	14	55	62	79	84	50	76	44	82	123
Per item (content density)	1.6	1.8	1.4	1.5	3.9	1.6	2.8	9.2	1.4	1.3	1.4	1.3	1.1	1.5	1.3	2.3
Concepts linked to ICF component	866 (88)	410 (93)	456 (85)	44 (81)	39 (91)	205 (96)	13 (93)	50 (91)	59 (95)	71 (90)	79 (94)	48 (96)	52 (68)	41 (93)	64 (78)	101 (82)
Body Function	233 (24)	91 (21)	142 (26)	14 (26)		42 (20)	2 (14)	8 (15)	25 (40)	24 (30)	27 (32)	15 (30)	8 (11)	7 (16)	35 (43)	26 (21)
Activity and participation	586 (60)	284 (64)	302 (56)	30 (56)	26 (60)	147 (69)	11 (79)	39 (71)	31 (50)	47 (59)	52 (62)	33 (66)	36 (47)	30 (68)	29 (35)	75 (61)
Environment	47 (5)	35 (8)	12 (2)		13 (30)	16 (8)		3 (5)	3 (5)				8 (11)	4 (9)		
Concepts linked to the ICF at																
1st level	39 (4)	18 (4)	21 (4)		6 (14)	2 (1)	3 (21)	3 (5)	4 (6)	3 (4)	3 (4)	3 (6)	4 (5)	1 (2)	4 (5)	3 (2)
2nd level	429 (44)	201 (46)	228 (42)	22 (41)	17 (40)	96 (45)	10 (71)	27 (49)	29 (47)	22 (28)	42 (50)	21 (42)	30 (39)	20 (45)	35 (43)	58 (47)
3rd level	385 (39)	190 (43)	195 (36)	22 (41)	16 (37)	106 (50)		20 (36)	26 (42)	42 (53)	33 (39)	23 (46)	18 (24)	20 (45)	22 (27)	37 (30)
4th level	13 (1)	1 (0)	12 (2)			1 (0)				4 (5)	1 (1)	1 (2)			3 (4)	3 (2)
Concepts not linked to the ICF	113 (12)	31 (7)	82 (15)	10 (19)	4 (9)	8 (4)	1 (7)	5 (9)	3 (5)	8 (10)	5 (6)	2 (4)	24 (32)	3 (7)	18 (22)	22 (18)
Not covered	58 (6)	12 (3)	46 (9)	3 (6)	2 (5)	4 (2)			3 (5)	6 (8)	2 (2)	1 (2)	20 (26)	2 (5)	7 (9)	8 (7)
Not definable	55 (6)	19 (4)	36 (7)	7 (13)	2 (5)	4 (2)	1 (7)	5 (9)		2 (3)	3 (4)	1 (2)	4 (5)	1 (2)	11 (13)	14 (11)

* Percentages are calculated based on the total number of concepts for each instrument.

Table V: The number of different ICF categories employed to represent the instruments' concepts and their frequency distribution across the four ICF components and the ICF levels of hierarchy for the selected instruments, and summarised for the generic and stroke specific measures.

	ICF total N	Total n (%)*	Generic n (%)*	Specific n (%)*	SF-36 n (%)*	RNL n (%)*	SIP n (%)*	EQ-5D n (%)*	LHS n (%)*	NHP n (%)*	SIS n (%)*	SSQoL n (%)*	SAQoL-39 n (%)*	QLI-SV n (%)*	SA-SIP30 n (%)*	BOSS n (%)*	HS Quale n (%)*
Number of different ICF categories																	
Total	1454	200 (14)	150 (10)	150 (10)	23 (2)	29 (2)	104 (7)	12 (1)	47 (3)	38 (3)	56 (4)	52 (4)	39 (3)	25 (2)	32 (2)	35 (2)	69 (5)
Per concept (content diversity)		0.20	0.34	0.28	0.43	0.67	0.49	0.86	0.85	0.61	0.71	0.62	0.78	0.33	0.73	0.43	0.56
ICF categories per component																	
Body Function	493	55 (11)	34 (7)	41 (8)	3 (1)		26 (5)	2 (0)	8 (2)	10 (2)	14 (3)	17 (3)	12 (2)	4 (1)	4 (1)	11 (2)	17 (3)
Activity and Participation	393	126 (32)	99 (25)	103 (26)	20 (5)	23 (6)	68 (17)	10 (3)	36 (9)	25 (6)	42 (11)	35 (9)	27 (7)	17 (4)	25 (6)	24 (6)	52 (13)
Environment	258	19 (7)	17 (7)	6 (2)		6 (2)	10 (4)		3 (1)	3 (1)				4 (2)	3 (1)		
ICF categories per level of hierarchy																	
1st level	30	8 (27)	6 (20)	8 (27)		2 (7)	1 (3)	2 (7)	3 (10)	3 (10)	3 (10)	3 (10)	2 (7)	2 (7)	1 (3)	4 (13)	2 (7)
2nd level	362	78 (22)	66 (18)	57 (16)	9 (2)	15 (4)	45 (12)	10 (3)	26 (7)	13 (4)	13 (4)	24 (7)	16 (4)	14 (4)	15 (4)	14 (4)	34 (9)
3rd level	926	109 (12)	77 (8)	81 (9)	14 (2)	12 (1)	57 (6)		18 (2)	22 (2)	38 (4)	24 (3)	20 (2)	9 (1)	16 (2)	14 (2)	30 (3)
4th level	136	5 (4)	1 (1)	4 (3)			1 (1)				2 (1)	1 (1)	1 (1)			3 (2)	3 (2)

* Percentages are calculated based on the total number of existing ICF codes N for each ICF component and each ICF level of hierarchy.

Table VI: Final linkage of concepts from generic and stroke-specific patient-centred health status measures to the respective ICF categories in the Body Functions component. The figures in each cell depict the number of items within the instrument that contained concepts linked to the respective ICF category.

BODY FUNCTIONS ICF category	Total	Gen-eric	Spe-cific	SF-36	RNL	SIP	EQ-5D	LHS	NHP	SIS	SS QoL	SA QoL-39	QLI-SV	SA-SIP 30	BOSS	HS Quale
Chapter b1 Mental functions	**159**	**57**	**102**	**12**		**23**	**1**	**6**	**15**	**15**	**17**	**11**	**6**	**4**	**29**	**20**
b110 Consciousness functions	1	1						1								
b114 Orientation functions	1	1						1								
b1140 Orientation to time	3	2	1			1		1		1						
b1141 Orientation to place	2	2				1		1								
b11421 Orientation to others	1	1				1										
b117 Intellectual functions	2	1	1					1								1
b126 Temperament and personality functions	6		6								2	1				3
b1263 Psychic stability	2		2								2					
b1266 Confidence	1		1													1
b1300 Energy level	19	8	11	4		3			1		3	2	2		4	
b1301 Motivation	3		3								1	1				1
b1302 Appetite	2	1	1			1					1					
b134 Sleep functions	7	5	2			3			2						2	
b1341 Onset of sleep	3	3				1			2							
b1342 Maintenance of sleep	2	2				1			1							
b1343 Quality of sleep	1	1							1							
b1400 Sustaining attention	3		3							1	1				1	
b144 Memory functions	10	2	8			1		1			2	1			3	2
b1440 Short-term memory	1		1							1						
b1441 Long-term memory	1		1							1						
b1442 Retrieval of memory	1		1							1						
b1470 Psychomotor control	2	2				2										
b152 Emotional functions	53	17	36	8		2	1		6	5	4	2	2		16	7
b1521 Regulation of emotion	9	4	5			3			1			1		3		1
b1600 Pace of thought	1		1							1						
b164 Higher-level cognitive functions	4	2	2			2								1		1
b1641 Organization and planning	2	1	1			1										1
b167 Mental functions of language	2		2									2				
b16700 Reception of spoken language	3		3							1					1	1
b16701 Reception of written language	1		1												1	
b16710 Expression of spoken language	7		7							3	1	1			1	1
b1801 Body image	2		2										2			
b1802 Experience of time	1	1							1							
Chapter b2 Sensory functions and pain	**23**	**15**	**8**	**2**		**2**	**1**	**2**	**8**		**3**		**2**			**3**
b210 Seeing functions	4	1	3					1			2					1
b2101 Visual field functions	1		1								1					
b230 Hearing functions	1	1						1								
b280 Sensation of pain	16	13	3	2		2	1		8				2			1
b28010 Pain in head and neck	1		1													1
Chapter b3 Voice and speech functions	**11**	**4**	**7**			**4**					**3**	**1**		**3**		
b310 Voice functions	1	1				1										
b320 Articulation functions	7	2	5			2					2	1		2		
b330 Fluency and rhythm of speech functions	3	1	2			1					1			1		
Chapter b4 Functions of the cardiovascular, haematological, immunological and respiratory systems	**10**	**5**	**5**			**3**			**2**		**2**	**2**				**1**
b4550 General physical endurance	5	3	2			3					1	1				
b4552 Fatiguability	5	2	3						2		1	1				1
Chapter b5 Functions of the digestive, metabolic and endocrine systems	**9**	**2**	**7**			**2**				**1**	**1**				**5**	
b510 Ingestion functions	1	1				1										
b5102 Chewing	2		2												2	
b5105 Swallowing	4		4								1				3	
b5253 Faecal continence	2	1	1			1				1						
Chapter b6 Genitourinary and reproductive functions	**4**	**2**	**2**			**2**				**1**						**1**
b6202 Urinary continence	2	1	1			1				1						
b640 Sexual functions	2	1	1			1										1
Chapter b7 Neuromusculoskeletal and movement-related functions	**17**	**6**	**11**			**6**				**7**	**1**	**1**			**1**	**1**
b730 Muscle power functions	1		1													1
b7300 Power of isolated muscles and muscle groups	2		2							2						
b7301 Power of muscles of one limb	2		2							2						
b755 Involuntary movement reaction functions	8	2	6			2				3	1	1			1	
b760 Control of voluntary movement functions	3	3				3										
b770 Gait pattern functions	1	1				1										

Table VII: Final linkage of concepts from generic and stroke-specific patient-centred health status measures to the respective ICF categories in the Activity and Participation component. The figures in each cell depict the number of items within the instrument that contained concepts linked to the respective ICF category.

ACTIVITY AND PARTICIPATION ICF category		Total	Gen-eric	Spe-cific	SF-36	RNL	SIP	EQ-5D	LHS	NHP	SIS	SS QoL	SA QoL -39	QLI -SV	SA -SIP 30	BOSS	HS Quale
Chapter d1 Learning and applying knowledge		**35**	**11**	**24**		**1**	**8**		**2**		**1**	**4**	**2**		**2**	**4**	**11**
d1	Learning and applying knowledge	2		2												1	1
d110	Watching	7	4	3		1	2		1			2					1
d160	Focusing attention	3	2	1			2								1		
d163	Thinking	9	2	7			1		1			1			1	2	3
d166	Reading	1		1													1
d170	Writing	5	2	3			2					1	1				1
d175	Solving problems	2		2							1					1	
d1751	Solving complex problems	1		1													1
d177	Making decisions	5	1	4			1						1				3
Chapter d2 General tasks and demands		**43**	**23**	**20**	**2**		**18**		**2**	**1**		**2**	**2**	**2**	**3**	**2**	**9**
d210	Undertaking a single task	2	1	1			1										1
d2102	Undertaking a single task independently	19	15	4			13		1	1			1		3		
d220	Undertaking multiple tasks	2		2													2
d2200	Carrying out multiple tasks	2	2				2										
d2201	Completing multiple tasks	2		2								1	1				
d2202	Undertaking multiple tasks independently	3		3													3
d230	Carrying out daily routine	9	3	6	2				1			1		2		1	2
d2301	Managing daily routine	1		1												1	
d2303	Managing one's own activity level	2	2				2										
d2400	Handling responsibilities	1		1													1
Chapter d3 Communication		**37**	**11**	**26**		**1**	**9**		**1**		**2**	**6**	**5**	**2**	**3**	**5**	**3**
d3	Communication	1		1												1	
d330	Speaking	18	4	14			3		1			4	3	2	1	2	2
d335	Producing nonverbal messages	1	1				1										
d3350	Producing body language	1	1				1										
d340	Producing messages in formal sign language	1	1				1										
d345	Writing messages	1		1												1	
d350	Conversation	2	1	1			1								1		
d3501	Sustaining a conversation	2	1	1			1								1		
d3504	Conversing with many people	2		2							1					1	
d360	Using communication devices and techniques	1	1			1											
d3600	Using telecommunication devices	4		4							1	1	1				1
d3601	Using writing machines	3	1	2			1					1	1				
Chapter d4 Mobility		**149**	**79**	**70**	**12**	**3**	**44**	**3**	**5**	**12**	**15**	**15**	**11**	**6**	**11**	**6**	**6**
d4	Mobility	7	4	3			2	2			1				1	1	
d410	Changing basic body position	4	4				2		1	1							
d4100	Lying down	5	3	2			3				1				1		
d4102	Kneeling	3	2	1	1		1									1	
d4103	Sitting	6	2	4			2				1	1	1		1		
d4104	Standing	2	1	1			1									1	
d4105	Bending	6	3	3	1		1			1		1	1			1	
d4150	Maintaining a lying position	2	2				2										
d4153	Maintaining a sitting position	4	3	1			2			1	1						
d4154	Maintaining a standing position	6	3	3			1			2	1	1	1				
d4201	Transferring oneself while lying	1	1				1										
d430	Lifting and carrying objects	5	2	3	2						1						2
d4300	Lifting	1	1		1												
d440	Fine hand use	2	1	1			1								1		
d4400	Picking up	1		1							1						
d4402	Manipulating	8	1	7			1					4	3				
d445	Hand and arm use	4	3	1			3					1					
d4451	Pushing	1	1		1												
d4452	Reaching	4	1	3						1		2	1				
d4453	Turning or twisting the hands or arms	3	1	2			1				2						
d450	Walking	19	8	11			4	1		3	2	3	2	2	1		1
d4500	Walking short distances	3	2	1	1		1				1						
d4501	Walking long distances	3	2	1	2												1
d4551	Climbing	16	9	7	2		5			2	2	1	1		1	1	1
d4552	Running	2	1	1	1												1
d460	Moving around in different locations	11	6	5			4		1	1					4	1	
d4600	Moving around within the home	5	5			1	3		1								
d4602	Moving around outside the home and other buildings	5	3	2		2			1					2			
d465	Moving around using equipment	5	1	4			1					1	1	2			
d470	Using transportation	1	1						1								
d4701	Using private motorized transportation	3	1	2			1				1				1		
d4702	Using public motorized transportation	1	1				1										
Chapter d5 Self-care		**57**	**32**	**25**	**2**	**4**	**17**	**3**	**5**	**1**	**6**	**7**	**3**		**3**	**2**	**4**
d5	Self-care	2	1	1				1				1					
d510	Washing oneself	2	2					1	1								

ACTIVITY AND PARTICIPATION ICF category		Total	Generic	Specific	SF-36	RNL	SIP	EQ-5D	LHS	NHP	SIS	SS QoL	SA QoL-39	QLI-SV	SA-SIP 30	BOSS	HS Quale
d5101	Washing whole body	11	5	6	1	1	3				1	1	1		1	1	1
d5202	Caring for hair	1	1						1								
d5204	Caring for toe nails	1		1							1						
d530	Toileting	6	3	3		1	1		1		1	1					1
d540	Dressing	14	9	5	1	1	4	1	1	1		1	1		1	1	1
d5400	Putting on clothes	1		1							1						
d5402	Putting on footwear	6	1	5			1				1	2	1		1		
d550	Eating	8	5	3			5				1	1					1
d560	Drinking	2	2				2										
d5701	Managing diet and fitness	2	2			1	1										
d5702	Maintaining one's health	1	1						1								
Chapter d6		**56**	**22**	**34**	**1**	**1**	**8**	**1**	**7**	**4**	**10**	**5**	**2**	**4**	**4**	**2**	**7**
Domestic life																	
d6	Domestic life	3	1	2						1				2			
d620	Acquisition of goods and services	1		1													1
d6200	Shopping	5	2	3			1		1		1				1		1
d630	Preparing meals	5	2	3					1	1		1	1			1	
d640	Doing housework	16	9	7	1	1	4	1	2		2	2	1			1	1
d6400	Washing and drying clothes and garments	4	2	2			1		1		1				1		
d6401	Cleaning cooking area and utensils	2		2							1						1
d6402	Cleaning living area	6	2	4			1			1	1				1		2
d6403	Using household appliances	1		1							1						
d6405	Disposing of garbage	1		1							1						
d650	Caring for household objects	3	2	1			1			1					1		
d6505	Taking care of plants, indoors and outdoors	3	1	2					1		1						1
d660	Assisting others	6	1	5					1		1	2		2			
Chapter d7		**78**	**38**	**40**	**5**	**5**	**17**	**1**	**5**	**5**	**3**	**7**	**3**	**10**	**3**	**7**	**7**
Interpersonal interactions and relationships																	
d7	Interpersonal interactions and relationships	10	4	6		1			1	2	1	1	1			1	2
d7100	Respect and warmth in relationships	2	1	1			1								1		
d7104	Social cues in relationships	2	1	1			1								1		
d720	Complex interpersonal interactions	1	1				1										
d7200	Forming relationships	3	1	2						1						1	1
d7202	Regulating behaviours within interactions	5	4	1			4								1		
d730	Relating with strangers	1	1						1								
d740	Formal relationships	1	1			1											
d7402	Relating with equals	1	1				1										
d7500	Informal relationships with friends	15	5	10	2	1	1		1		1	2	1	2		3	1
d7501	Informal relationships with neighbours	4	2	2	1				1					2			
d760	Family relationships	21	13	8	2	2	6	1	1	1	1	3	1			2	1
d7600	Parent-child relationships	5	2	3			2							2			1
d770	Intimate relationships	3		3										2			1
d7702	Sexual relationships	4	1	3						1		1		2			
Chapter d8		**50**	**24**	**26**	**3**	**3**	**9**	**2**	**6**	**1**	**2**	**1**		**8**			**15**
Major life areas																	
d820	School education	6	2	4		1		1						2			2
d830	Higher education	1		1													1
d8451	Maintaining a job	1		1													1
d850	Remunerative employment	24	13	11	3	1	6	1	1	1	1	1		4			5
d8501	Part-time employment	1		1													1
d8502	Full-time employment	3	2	1			2										1
d855	Non-remunerative employment	3	2	1		1			1		1						
d860	Basic economic transactions	2	1	1			1										1
d870	Economic self-sufficiency	2	1	1					1								1
d8700	Personal economic resources	7	3	4					3					2			2
Chapter d9		**81**	**44**	**37**	**5**	**8**	**17**	**1**	**6**	**7**	**8**	**5**	**5**	**4**	**1**	**1**	**13**
Community, social and civic life																	
d9	Community, social and civic life	5	1	4					1		1	1	2				
d910	Community life	3	2	1	1		1										1
d920	Recreation and leisure	21	11	10		1	6	1	1	2	2	2	1	2			3
d9200	Play	3	3			1	2										
d9201	Sports	7	5	2	2	1			1	1	1						1
d9202	Arts and culture	8	5	3		1	2		1	1	1						2
d9203	Crafts	3	2	1		1				1	1						
d9204	Hobbies	7	4	3		1	1		1	1		1	1				1
d9205	Socializing	18	11	7	2	2	5		1	1		1	1		1	1	3
d930	Religion and spirituality	2		2										2			
d9300	Organized religion	2		2							1						1
d9301	Spirituality	2		2							1						1

Table VIII: Final linkage of concepts from generic and stroke-specific patient-centred health status measures to the respective ICF categories in the component Environment. The figures in each cell depict the number of items within the instrument that contained concepts linked to the respective ICF category.

ENVIRONMENT ICF category		Total	Generic	Specific	SF-36	RNL	SIP	EQ-5D	LHS	NHP	SIS	SS QoL	SA QoL -39	QLI -SV	SA -SIP 30	BOSS	HS Quale
Chapter e1 **Products and technology**		**30**	**25**	**5**		**8**	**14**		**1**	**2**				**2**	**3**		
e1100	Food	2	2				2										
e1101	Drugs	1	1							1							
e115	Products and technology for personal use in daily living	2	2			2											
e1150	General products and technology for personal use in daily living	1	1				1										
e1151	Assistive products and technology for personal use in daily living	2	2				2										
e120	Products and technology for personal indoor and outdoor mobility and transportation	7	5	2			3		1	1					2		
e1201	Assistive products and technology for personal indoor and outdoor mobility and transportation	6	6			3	3										
e130	Products and technology for education	1	1			1											
e135	Products and technology for employment	1	1			1											
e1350	General products and technology for employment	1	1				1										
e1351	Assistive products and technology for employment	1	1				1										
e140	Products and technology for culture, recreation and sport	1	1			1											
e155	Design, construction and building products and technology of buildings for private use	4	1	3			1							2	1		
Chapter e2 **Natural environment and human-made changes to environment**		**2**	**1**	**1**			**1**								**1**		
e240	Light	2	1	1			1								1		
Chapter e3 **Support and relationships**		**12**	**8**	**4**		**5**	**1**		**1**	**1**				**4**			
e3	Support and relationships	9	7	2		5			1	1				2			
e310	Immediate family	2		2										2			
e325	Acquaintances, peers, colleagues, neighbours and community members	1	1				1										
Chapter e5 **Services, systems and policies**		**3**	**1**	**2**					**1**					**2**			
e5700	Social security services	1	1						1								
e5800	Health services	2		2										2			

IV. SELECTING HEALTH STATUS MEASURES BASED ON CONTENT VALIDITY: Comparison of stroke-specific health status measures with the *Comprehensive ICF Core Set for Stroke*

IV.1. Specific aims

The aim of the second study is demonstrate how the use of the ICF as a fundamental reference can be a useful approach to select health status measures according to their content validity. The specific aims are (1) to examine the content validity of the selected stroke-specific health status measures by comparing them with the *Comprehensive ICF Core Set for Stroke*, and (2) to discuss the selection of measures based on their coverage of the ICF Core Set.

IV.2. Methods

Seven stroke-specific patient-centered health status measures are involved in the current analyses and compared to the *Comprehensive ICF Core Set for Stroke*: the Stroke Impact Scale[48,87] (SIS), the Stroke-specific Quality of Life Scale (SSQoL), the Stroke and Aphasia Quality of Life Scale (SAQoL-39), the Quality of Life Index - Stroke Version (QLI-SV), the Stroke-Adapted Sickness Impact Profile-30 (SA-SIP-30), the Burden of Stroke Scale[42,47] (BOSS), and the Quality of Life Instrument for Young Hemorrhagic Stroke Patients (HSQuale). Within the previous study, the measures contents have been translated into the taxonomy of the ICF applying the linking method. The comparison of the measures with the *Comprehensive ICF Core Set for Stroke* is conducted based on these linking results. As the seven measures as well as the ICF Core Set for Stroke are represented in the form of lists of ICF categories, they can be cross-tabulated and contrasted against each other.

The *Comprehensive ICF Core Set for Stroke* includes a total of 130 categories at the 2nd level of the ICF classification. However, the results of the measures' linking to the ICF also contain categories at the more specific 3rd and 4th levels. Therefore, to accomplish the content comparison, the measures' linking results are cumulated at the 2nd level of the ICF. The categories of the ICF are arranged in a stem/branch/leaf scheme. Consequently, a more specific lower-level category shares the attributes of the more global higher-level category to which it belongs, i.e., the use of a lower-level category automatically implies that the higher-level category is applicable. Therefore, for the purpose of the current analyses the more specific 3rd and 4th level ICF categories used to link the concepts within the measures can be recoded to yield the linking results in form of a list of ICF categories at the 2nd level. For example, the item from the SA-SIP30 "I am doing fewer social activities with groups of people" has been linked to the 3rd level ICF category *d9205 Socializing*. For the current analyses instead of the 3rd level category the overlying 2nd level category *d920 Recreation and leisure* is coded. For each of the 2nd level ICF categories the number of different 3rd and 4th level sub-categories that have been recoded in this way is documented.

To compare the contents of the stroke-specific patient-centered health status measures with the *Comprehensive ICF Core Set for Stroke* frequency analyses are conducted in two ways. First, for each category of the ICF Core Set and for each instrument the dichotomous information whether the measure covers that category or not is analyzed to indicate the instruments' bandwidth of content coverage regarding the ICF Core Set. The number of 2nd level ICF Core Set categories covered by the instruments is summed up. Also, for each instrument the percentage of the ICF Core Set's categories covered is given. The results are presented overall as well as for each component of the ICF (i.e., for Body Functions, Body Structures, Activity and Participation, Environmental Factors). ICF Core Set categories not covered by the instruments, as well as ICF categories included in the instruments but not part of the ICF Core Set are listed.

Second, for each of the 2^{nd} level categories of the *Comprehensive ICF Core Set for Stroke*, the number of more detailed 3^{rd} and 4^{th} level categories covering the according areas of functioning in the different measures is considered to indicate their specificity of content coverage. The number of categories at the more detailed 3^{rd} and 4^{th} levels within the instruments is summed up within each 2^{nd} level category. In addition, the number of specific 3^{rd} and 4^{th} level categories is also shown for each instrument overall and for the different components of the ICF.

IV.3. Results

Table IX shows the ICF categories shared by the *Comprehensive ICF Core Set for Stroke* and the seven examined stroke-specific patient-centered health status measures. Each ICF category's coverage and the number of 3^{rd} and 4^{th} level categories used are displayed. Table X summarizes the number of ICF Core Set categories covered by the instruments, the percentage of ICF Core Set coverage, as well as the number of the additional 3^{rd} and 4^{th} level categories covered. 67 (52%) out of the 130 categories of the ICF Core Set are covered by at least one of the examined instruments. 41 categories from the component Activity and Participation representing 80% of the ICF Core Set's categories from this component are covered by at least one of the measures. In the component of Body Functions 22 categories (54%) and in the component of Environmental Factors 4 categories (12%) are addressed by at least one of the instruments.

63 or 48% of the ICF Core Set's categories are not covered by any of the instruments. These categories are listed in Table XI. In contrast, 11 ICF categories

are addressed by at least one of the examined instruments but are not part of the *Comprehensive ICF Core Set for Stroke*. These categories are shown in table XII.

The single measures cover in total between 29% (38 categories, HSQuale) and 14% (18 categories, QLI-SV) of the *Comprehensive ICF Core Set for Stroke*. With the HSQuale, the SSQoL (35 categories, 27%) and the SIS (32 categories, 25%) represent the top three instruments with the largest coverage of the relevant areas of stroke patients' functioning as represented by the ICF Core Set.

All instruments cover Activity and Participation as well as Body Functions, the former being consistently regarded to a larger proportion in all of the examined measures. On the other hand, only two instruments address Environmental Factors, namely the SA-SIP-30 (2 categories, 6%) and the QLI-SV (3 categories, 9%), although the *Comprehensive ICF Core Set for Stroke* contains 33 Environmental Factors. No categories of the ICF component Body Structures are contained in the examined instruments, while the ICF Core Set includes 5 categories from this component.

Overall, 31 categories of the ICF Core Set are measured at the more detailed 3rd and 4th levels by at least one of the seven instruments. In contrast, 36 categories are addressed by the instruments at the 2nd level only. The single measures include content concepts that have been linked to 19 (HS-Quale) to 2 (QLI-SV) different ICF categories at the more detailed 3rd and 4th levels of the classification. Five categories are measured in-depth at the more specific levels of the ICF by at least 3 of the instruments: *d410 Changing basic body position, b167 Mental functions of language, d540 Dressing, d920 Recreation and leisure, d640 Doing housework.* Beyond the HSQuale, again the SIS (16 categories) and the SS-QoL (11 categories) are among the top three instruments including more specific contents within the 2nd level categories of the *Comprehensive ICF Core Set for Stroke*.

IV.4. Discussion

In the current study, seven stroke-specific patient-centered health status measures have been compared with the *Comprehensive ICF Core Set for Stroke*. This ICF Core Set based on the universal common language of the ICF can serve as a standard to determine what to measure in stroke. Therefore, the comparison of instruments against this standard can be used to characterize and compare measures' content validity. Using this method of comparison between the *Comprehensive ICF Core Set for Stroke* and the instruments measuring stroke-specific health status two aspects of content validity are considered: bandwidth and specificity of content coverage.

The examined instruments only cover one third to one fifth of the *Comprehensive ICF Core Set for Stroke*. Categories from the component Activity and Participation are most frequently covered followed by Body Functions. A high number of ICF Core Set categories is not addressed. No Body Structures and only few Environmental Factors are measured by the examined instruments.

The current analyses, by summing up the number of categories covered, rely on the simplified assumption that the different ICF categories have equal relevance. However, several of the categories not covered by the instruments are obviously of low relevance in a subpopulation of stroke patients who are able to complete a self-report measure, e.g. *b110 Consciousness functions*, *d325 Communicating with – receiving – written messages*, *d310 Communicating with – receiving – spoken messages*. Patients with clouded consciousness or receptive aphasia would not be able to fill in a self-report measure on health status, thus according categories are obviously not relevant to be covered in this specific type of instruments.

One further reason, why several categories of the ICF Core Set are not covered by the instruments can be given. Categories of the ICF can be regarded as related to each other across the different components in accordance to the basic biopsychosocial model, for example certain impairments of Body Functions and Structures may lead to specific difficulties in Activities and Participation. The examined patient-centered measures seem to emphasize the perspective of Activity and Participation as outcome, over and above Body Function and Structure as precondition, which are therefore less frequently covered. For example, Body Function and Structure categories from the ICF Core Set, like *b176 Mental functions of sequencing complex movements*, *b735 Muscle tone functions*, *b710 Mobility of joint functions*, *s730/ s750 Structure of upper/ lower extremity* are not included in the instruments. However, chapter *d4 Mobility* within the Activity and Participation component is represented at length in the instruments by several categories like *d450 Walking*, *d410 Changing basic body position*, *d445 Hand and arm use*, etc. The emphasis on the Activity and Participation perspective when a connection of the categories across different ICF components is apparent counterbalances the first impression, that the measures miss a high number of relevant areas of functioning and thus might be compromised with respect to their content validity.

However, Environmental Factors are not covered at all by five out of the seven examined measures and the two remaining instruments only cover four different ICF categories, while the *Comprehensive ICF Core Set for Stroke* comprises 33 categories within this component. Environmental Factors play an important role for patients' functioning in stroke as facilitators or barriers.[154] Especially the significance of the social environment and of family support,[155] of medication,[156,157,158,159] and of different types of health services[160,161,162] is well established as they may improve functioning following stroke. Although the environment influences patients' health status, the measures examined here have not been designed to capture the explanatory factors which work together to produce health status. Other instruments would be utilized to measure patients' environment[163,164,165,166] or the influence of the environment on functioning.[167]

Examining the measures specificity of content coverage revealed that about half of the categories shared between the *Comprehensive ICF Core Set for Stroke* and the instruments are measured at a more in-depth level. The ICF's 2nd level categories differ as to the number of more specific categories they are containing. Thus, multi-faceted 2nd level categories can be distinguished in contrast from 2nd level categories representing a single unsplit conceptual unit. For example, the 2nd level ICF category *d920 Recreation and leisure* includes 6 more specific categories at the 3rd level, namely *d9200 Play*, *d9201 Sports*, *d9202 Arts and culture*, *d9203 Crafts*, *d9204 Hobbies*, and *d9205 Socializing*. On the other hand, the 2nd level ICF category *d330 Speaking*, for example, does not include any more detailed sub-categories at the 3rd and 4th levels.

In the current analyses, those ICF Core Set categories that are covered at a more detailed level by at least one of the examined stroke-specific health status measures are such categories of the classification that are to a higher degree multi-faceted. In comparison, those categories which are covered by the instruments only at the more global 2nd level are more frequently such ICF categories not including any more detailed sub-categories at the 3rd and 4th levels. In this way, as expected, by the linking method the structure of the ICF is reflected, as more concepts of the measures fall into the broader multi-faceted categories. Therefore, if broad ICF categories are covered by using several sub-categories this confirms a more comprehensive content coverage, and thus, indicates better content validity.

Accounting for both, the bandwidth as well as the specificity of content coverage, a pool of instruments can be selected and shortlisted for further consideration. The three instruments with the highest level of content validity are the HSQuale, the SSQoL and the SIS. The HSQuale is according to bandwidth as well as specificity of content coverage the most comprehensive of the examined measures. However, it is developed for the specific population of young patients with hemorrhagic stroke. In contrast, the SSQoL and SIS are both targeted at stroke survivors without restriction of etiology or age. Although the SSQoL proved to be the measure with a somewhat larger bandwidth of content coverage, the SIS ranks before the SSQoL regarding specificity. Thus, further features of the instruments need to be considered to arrive at a final choice.

The usefulness of the ICF based content examination for the selection of appropriate instruments shows to practical advantage especially when further steps are considered. First, the purpose of the data collection can also be formulated using the category system of the ICF. In this way, the purpose of the investigation and the assessment instruments could be matched to each other one to one. For example, special methods for the linking of interventions, as well as a study applying the linking to nursing interventions are available.[168] Second, the proliferation of studies presenting the linking results of further health status measures would lessen the efforts of conducting the linking procedure in future. Such studies have been published on specific measures in various health conditions[169,170,171,172] and for a number of generic quality of life instruments.[173] Third, the selection of appropriate instruments using the ICF based content examination approach may be further facilitated and eased with the adoption of a simple database as it has been used for the current study.

Content validity, defined as the extent to which an instrument measures the full scope of the concept to be measured[174] is established during the process of instrument development[175] by involving patient, health professional, expert or caregiver input using qualitative methods, like individual or focus group interviews, or semi-quantitative methods, like expert rating of content coverage.[176] While all of the examined measures have involved during their development process qualitative input from patients and health professionals, the item generation and selection followed different rationales and arrived at different contents covered. Still, the examination and comparison of content coverage across the different measures developed applying different techniques is now possible by the use of the ICF Core Sets as common reference along with the linking method.

Several current guidelines on post-stroke rehabilitation and care deal with the question what to measure in stoke. They refer to different levels of abstraction, have a different main focus and include a variety of areas affected by stroke without being fully consistent or comparable.[177,178,179,180] In addition, these guidelines do not contain recommendations on specific instruments to be used. Although the current analyses are limited to patient-centered measures only, in further studies also other types of measures could be involved. While the examined patient-centered measures seem to focus mainly on Activity and Participation, it can be assumed that standardized rating scales or performance tests frequently used in stroke,[30,31,32] may cover further parts of the ICF Core Set, which are not covered by the measures selected for the current examination. In future studies, based on the *Comprehensive ICF Core Set for Stroke* as a common reference a combination of different health status measures can be identified that ensures the broad yet efficient coverage of those areas of stroke patients' functioning, which have been regarded most relevant and are part of the

ICF Core Set. In this way, scientifically founded recommendations on instruments to be used could be developed in future.

The current analyses revealed a small number of ICF categories that are addressed by the instruments but are not part in the list of relevant categories of the *Comprehensive ICF Core Sets for Stroke*. These categories should be considered within the validation of the ICF Core Sets as potential candidate categories to be amended. Further investigation is needed to clarify and agree upon their relevance.

The comparison of the seven selected instruments' bandwidth and specificity of content coverage have been conducted to indicate their content validity, i.e. their representativity regarding the ***constructs to be measured***. An important limitation of the current study refers to the fact that the instruments examined rely on primarily different conceptualizations of patient-centered health status. They include (health-related) quality of life, subjective well-being,[181] life satisfaction,[182] sickness,[183] and disease consequences according to the ICIDH-Model of impairment, disability, and handicap, as well as the concept of functioning according to the ICF. Measures based on various conceptualizations have been mapped to and compared with the ICF Core Set derived from the WHO's biopsychosocial model and conceptualization of functioning. However, the relationship between the different concepts and the ICF is not clarified, yet. Therefore, content validity and comprehensiveness of content coverage are evaluated here against a reference which is to a certain degree alien to the primary constructs of some of the examined measures. Thus, for the careful interpretation of the results in terms of content validity this constraint needs to be considered. It seems obvious, that an instrument developed to measure health status in the sense of life satisfaction will rarely cover Body Functions and Structures, major components of functioning and health within the framework of the ICF. However, the ICF and also the ICF Core Set are highly comprehensive and embrace different perspectives on health. Therefore, they seem to be well suited to serve as external reference to map onto a variety of different conceptual issues.

The examination and comparison of patient-centered health status measures' content validity based on the *Comprehensive ICF Core Set for Stroke* and the linking method can serve as a first step of selecting a measure. The results of the current study established a tentative rank order of the examined instruments' content validity, with the HSQuale, the SIS, and the SS-QoL at the highest ranks. However, further features of the measures have to be considered. Especially, their psychometric properties have to be carefully examined to accomplish the well-founded choice of an appropriate measure to assess stroke related health status. How to do this based on modern probabilistic test theory will be shown in the next chapter.

Table IX: ICF categories shared by the Comprehensive ICF Core Set for Stroke and the seven examined stroke-specific patient-centered health status measures. The figures in brackets indicate the number of different 3rd and 4th level ICF categories that have been cumulated to yield the information at the 2nd level.

ICF Code 2nd level	ICF category title	ICF Core Set for Stroke	HS-Quale	SS-QoL	SIS	SA-QoL-39	BOSS	SA-SIP-30	QLI-SV
	Body Functions								
b114	Orientation functions	1			1				
b117	Intellectual functions	1	1						
b126	Temperament and personality functions	1	1(+1)	1(+1)		1			
b130	Energy and drive functions	1	1	1(+2)		1(+1)	1		1
b134	Sleep functions	1					1		
b140	Attention functions	1		1	1		1		
b144	Memory functions	1	1	1	1(+2)	1	1		
b152	Emotional functions	1	1(+1)	1	1	1(+1)	1	1	1
b164	Higher-level cognitive functions	1	1(+1)					1	
b167	Mental functions of language	1	1(+1)	1	1(+1)	1(+1)	1(+2)		
b180	Experience of self and time functions	1							1
b210	Seeing functions	1	1	1(+1)					
b280	Sensation of pain	1	1(+1)						1
b320	Articulation functions	1		1		1		1	
b330	Fluency and rhythm of speech functions	1		1				1	
b455	Exercise tolerance functions	1	1	1(+1)		1(+1)			
b510	Ingestion functions	1		1			1(+1)		
b525	Defecation functions	1			1				
b620	Urination functions	1			1				
b640	Sexual functions	1	1						
b730	Muscle power functions	1	1		1(+1)				
b755	Involuntary movement reaction functions	1		1	1	1	1		
	Activity and Participation								
d160	Focusing attention	1						1	
d166	Reading	1	1						
d170	Writing	1	1	1		1			
d175	Solving problems	1	1		1		1		
d210	Undertaking a single task	1				1		1(+2)	
d220	Undertaking multiple tasks	1	1(+1)	1		1			
d230	Carrying out daily routine	1	1	1			1(+1)		1
d240	Handling stress and other psychological demands	1	1						
d330	Speaking	1	1	1		1	1	1	1
d345	Writing messages	1					1		
d350	Conversation	1			1		1	1(+1)	
d360	Using communication devices and techniques	1	1	1(+1)	1	1(+1)			

ICF Code 2nd level	ICF category title	ICF Core Set for Stroke	HS-Quale	SS-QoL	SIS	SA-QoL-39	BOSS	SA-SIP-30	QLI-SV
d410	Changing basic body position	1		1(+1)	1(+1)	1(+1)	1(+2)	1(+1)	
d415	Maintaining a body position	1		1	1(+1)	1			
d430	Lifting and carrying objects	1	1		1				
d440	Fine hand use	1		1	1	1		1	
d445	Hand and arm use	1		1(+1)	1	1			
d450	Walking	1	1(+1)	1	1(+1)	1		1	1
d455	Moving around	1	1(+1)	1	1	1	1	1	
d460	Moving around in different locations	1					1	1	1
d465	Moving around using equipment	1		1		1			1
d470	Using transportation	1			1			1	
d510	Washing oneself	1	1	1	1	1	1	1	
d520	Caring for body parts	1			1				
d530	Toileting	1	1	1	1				
d540	Dressing	1	1	1(+1)	1(+1)	1(+1)	1	1(+1)	
d550	Eating	1	1	1	1				
d620	Acquisition of goods and services	1	1(+1)		1			1	
d630	Preparing meals	1		1		1	1		
d640	Doing housework	1	1(+2)	1	1(+5)	1	1	1(+1)	
d710	Basic interpersonal interactions	1						1(+1)	
d750	Informal social relationships	1	1	1	1	1	1		1(+1)
d760	Family relationships	1	1(+1)	1	1	1	1		1
d770	Intimate relationships	1	1	1					1(+1)
d845	Acquiring, keeping and terminating a job	1	1						
d850	Remunerative employment	1	1(+2)	1	1				1
d855	Non-remunerative employment	1			1				
d860	Basic economic transactions	1	1						
d870	Economic self-sufficiency	1	1(+1)						1
d910	Community life	1	1						
d920	Recreation and leisure	1	1(+4)	1(+2)	1(+3)	1(+2)	1	1	1
	Environmental Factors								
e120	Products and technology for personal indoor and outdoor mobility and transportation	1						1	
e155	Design, construction and building products and technology of buildings for private use	1						1	1
e310	Immediate family	1							1
e580	Health services, systems and policies	1							1

Table X: Number and percentage of the categories of the Comprehensive ICF Core Set for Stroke covered by the instruments and the number of the additional 3rd and 4th level categories addressed by the instruments is shown overall as well as for the different components of the ICF.

ICF Code 2nd level	ICF category title	Total	HS-Quale	SS-QoL	SIS	SA-QoL-39	BOSS	SA-SIP-30	QLI-SV
	Overall								
	Number of ICF Core Set categories covered	67	38	35	32	27	23	22	18
	Portion of ICF Core Set categories covered (n=130)	52%	29%	27%	25%	21%	18%	17%	14%
	Number of 3rd and 4th level categories covered		19	11	16	9	6	7	2
	Body Functions								
	Number of ICF Core Set categories covered	22	12	12	9	8	8	4	4
	Portion of ICF Core Set categories covered (n=41)	54%	29%	29%	22%	20%	20%	10%	10%
	Number of 3rd and 4th level categories covered		5	5	4	4	3	0	0
	Activity and Participation								
	Number of ICF Core Set categories covered	41	26	23	23	19	15	16	11
	Portion of ICF Core Set categories covered (n=51)	80%	51%	45%	45%	37%	29%	31%	22%
	Number of 3rd and 4th level categories covered		14	6	12	5	3	7	2
	Environmental Factors								
	Number of ICF Core Set categories covered	4	0	0	0	0	0	2	3
	Portion of ICF Core Set categories covered (n=33)	12%	0%	0%	0%	0%	0%	6%	9%
	Number of 3rd and 4th level categories covered		0	0	0	0	0	0	0

Table XI: Listing of the categories of the Comprehensive ICF Core Set for Stroke not covered by the seven examined stroke-specific patient-centred health status measures.

ICF Code 2nd Level	ICF category title
Body Functions	
b110	Consciousness functions
b156	Perceptual functions
b172	Calculation functions
b176	Mental function of sequencing complex movements
b215	Functions of structures adjoining the eye
b260	Proprioceptive function
b265	Touch function
b270	Sensory functions related to temperature and other stimuli
b310	Voice functions
b410	Heart functions
b415	Blood vessel functions
b420	Blood pressure functions
b710	Mobility of joint functions
b715	Stability of joint functions
b735	Muscle tone functions
b740	Muscle endurance functions
b750	Motor reflex functions
b760	Control of voluntary movement functions
b770	Gait pattern functions
Body Structures	
s110	Structure of brain
s410	Structure of cardiovascular system
s720	Structure of shoulder region
s730	Structure of upper extremity
s750	Structure of lower extremity
Activity and Participation	
d115	Listening
d155	Acquiring skills
d172	Calculating
d310	Communicating with - receiving - spoken messages
d315	Communicating with - receiving - nonverbal messages
d325	Communicating with - receiving - written messages

ICF Code 2nd Level	ICF category title
d335	Producing nonverbal messages
d420	Transferring oneself
d475	Driving
d570	Looking after one's health
Environmental Factors	
e110	Products or substances for personal consumption
e115	Products and technology for personal use in daily living
e125	Products and technology for communication
e135	Products and technology for employment
e150	Design, construction and building products and technology of buildings for public use
e165	Assets
e210	Physical geography
e315	Extended family
e320	Friends
e325	Acquaintances, peers, colleagues, neighbours and community members
e340	Personal care providers and personal assistants
e355	Health professionals
e360	Other professionals
e410	Individual attitudes of immediate family members
e420	Individual attitudes of friends
e425	Individual attitudes of acquaintances, peers, colleagues, neighbours and community members
e440	Individual attitudes of personal care providers and personal assistants
e450	Individual attitudes of health professionals
e455	Individual attitudes of health-related professionals
e460	Societal attitudes
e515	Architecture and construction services, systems and policies
e525	Housing services, systems and policies
e535	Communication services, systems and policies
e540	Transportation services, systems and policies
e550	Legal services, systems and policies
e555	Associations and organizational services, systems and policies
e570	Social security services, systems and policies
e575	General social support services, systems and policies
e590	Labour and employment services, systems and policies

Table XII: ICF Categories covered by at least one of the seven examined stroke-specific patient-centered health status measures but not included in the Comprehensive ICF Core Set for Stroke.

ICF Code 2nd level	ICF category title	HS-Quale	SS-QoL	SIS	SA-QoL-39	BOSS	SA-SIP-30	QLI-SV
	Body Functions							
b160	Thought functions			1				
	Activity and Participation							
d110	Watching	1	1					
d163	Thinking	1	1			1	1	
d177	Making decisions	1			1			
d650	Caring for household objects	1		1			1	
d660	Assisting others		1	1				1
d720	Complex interpersonal interactions	1				1	1	
d820	School education	1						1
d830	Higher education	1						
d930	Religion and spirituality	1		1				1
	Environmental Factors							
e240	Light						1	

V. APPLYING THE RASCH METHOD:

Evaluation of the Stroke Impact Scale using Rasch Analyses

V.1. Specific aims

The objective of the third study is to show, how the psychometric features of health status measures can be examined based on modern test theory and Rasch based methods. The psychometric properties of the Stroke Impact Scale (SIS), which was one of the previously examined and selected instruments, is evaluated in a German population.

The specific aims are to examine (1) the unidimensionality of the single domains and item fit, (2) the structure of the response scale, (3) the targeting of the domains, (4) reliability, (5) differential item functioning for relevant patient groups, and (6) to compare the fit results of this study with the Rasch analysis results of the SIS 2.0 in a North American sample, which led to the creation of the most current version of the SIS, the SIS 3.0.

V.2. Methods

V.2.1. The Stroke Impact Scale 2.0

The Stroke Impact Scale (SIS) is an established and well-examined stroke-specific QoL measure.[48,145,146,147,148,184] Although recently developed, the SIS has already been in use for clinical documentation and quality management purposes,[185,186] as well as within a large scale randomized controlled drug trial.[187] The SIS has been developed in a multi-stage process using input from patient, caregiver and professional perspectives following most current instrument development standards.[87] The SIS development was guided by the conceptual framework of the International Classification of Impairments, Disabilities and Handicaps (ICIDH) model to capture the full range of stroke outcomes. The most current version of the Stroke Impact Scale SIS 3.0, has been created by item reduction using Rasch analysis.

The Rasch analysis conducted by Duncan et al (2003) that resulted in the SIS 3.0 was performed on SIS 2.0 data from a sample of 696 mild to moderately affected stroke patients from the United States and Canada. They examined the reliability of the SIS domains, the unidimensionality of the domain constructs, the targeting, and the difficulty hierarchy of the items. However, they report on the functioning of the response category scales only for one domain and have not considered differential item functioning or item bias with regard to relevant subgroups of the sample, which are both valuable techniques within the Rasch approach to examine validity.

The SIS 2.0 version is the first stroke-specific quality of life measure that has been translated into 14 languages and is being psychometrically evaluated in international research efforts. In Germany, Petersen et al (2001) report on the translation process and the psychometric evaluation of the SIS 2.0 applying the traditional test theory (TTT) paradigm.[188] In a sample of 137 stroke patients the internal consistency reliability, criteria-related and convergent validity, and scale fit of the German version of the SIS was studied. Currently, psychometric evaluation of the German SIS using Rasch methods has not been conducted.

The Stroke Impact Scale covers 8 domains that represent distinct aspects of stroke-related QoL: Strength, Memory/ Thinking, Emotion, Communication, ADL/ IADL, Mobility, Hand Function, and Participation. SIS Version 2.0 consists of 64 items rated along a 5 point scale. Aggregate scores for each domain are generated by a specific algorithm and range from 0 to 100 with higher scores meaning higher levels of quality of life. No single summary score is generated. However, the 4 physical domains (Strength, ADL/ IADL, Mobility, Hand Function) can be summed up to create one combined score. The SIS 2.0 proved to be a reliable, valid and sensitive measure of stroke-related QoL. In German speaking countries so far no other stroke-specific QoL measure is available that has been psychometrically tested. Table XIII shows item examples from each domain of the SIS and the respective response categories. Appendix 3 contains the German version of the Stroke Impact Scale 2.0 as it was applied in the study.

V.2.2. Study design

The psychometric evaluation of the SIS is conducted using cross-sectional data collected from stroke patients in Germany within an ongoing multicentric international study. The parent study is carried out in cooperation with the WHO and is part of efforts regarding the application of the *ICF – International Classification of Functioning, Disability and Health*.[49,65] The study protocol and the informed consent forms were approved by the responsible Ethic Committees in each involved world region. In Germany, 32 study centers providing acute or rehabilitation services to stroke patients are currently participating in the data collection.

V.2.3. Participants

Stroke patients are eligible when they gave informed consent to participate and signed a consent form, were older than 18 years, have suffered from an acute cerebrovascular disorder, their main diagnosis was coded with the ICD-10 codes I60-

I64, and an attending health professional gave positive judgment of the ability of the patient to complete alone or with assistance of a proxy self-report assessments. The patient was excluded, when he or she had a surgery's wound that was not completely healed at the time of the data collection.

Data were collected using a battery of patient-centered health status measures including the SIS 2.0. The data collection also included demographic information on each patient, stroke diagnosis according to ICD-10, date of stroke diagnosis, functional status prior to stroke, current health care setting (inpatient or outpatient), and Rankin Scale score.[189]

V.2.4. Analyses

Descriptive statistics to characterize the study population, and SIS scores are calculated using SAS software. Variables include stroke patients' gender, age, stroke etiology, time since stroke diagnosis, current health care setting, Rankin Scale Score, and functioning prior to stroke. For descriptive purposes SIS domain scores generated according to the standard scoring algorithm are reported. In contrast, for the following Rasch analyses, the raw scores are required and no imputation technique is applied. The estimation processes within the Rasch framework readily deal with missing values. Rasch analyses are conducted by the WINSTEPS software[190] applying the partial credit model.[191]

First, the ***unidimensionality*** of the single domains is examined. Unidimensionality refers to the idea that items should contribute to the measurement of only one single attribute at a time and should not be confounded by other attributes or dimensions. This idea is an essential aspect of construct validity. Unidimensionality is an inherent feature of the Rasch family of models, thus, by comparing the accord of the observed response data in a set of items to the expected values predicted by the Rasch model unidimensionality can be verified.[100,101,191,192] The ***fit*** of the items, i.e. their accord to the model, is examined using the infit mean square statistic, which is an information-weighted standardized residual accounting for the impact of unexpected outlier response patterns. The infit mean square has an expected value of 1.0. Items with an infit larger than 1.3 are considered misfitting.[193]

The ***structure of the response scale*** is studied based on the ordering of the threshold estimates for each item's response categories. The threshold parameters should take increasing values, as they represent successive transition points along the item response scale. With reversed thresholds it is apparent that the response scale does not function the way intended and the interpretation of scale scores has limited validity.[194]

Misfit of items and disordered category thresholds question the validity of the estimated parameters, thus the validity of further conclusions based on them. Therefore, the ***data are purified*** in two steps, namely collapsing of response categories and stepwise deletion of misfitting items. Collapsing is done by (1) systematically testing all options for creating a four-category or a three-category response scale and (2) selecting the ideal option for each domain to be applied to all items according to the following criteria: minimal change in the response scale, ordered thresholds, minimal number of misfitting items, highest reliability, and mean threshold distance between 1.0 to 5.0 logits. The stepwise item deletion starts with the most misfitting item per domain. After deleting the most misfitting item, the model is recalibrated and unidimensionality and item fit is rechecked. This procedure is repeated until no item shows model misfit.

Reliability is studied with the ***person reliability index***, which is analogous to the traditional test theory indices Kuder-Richardson Formula 21 or Cronbach's alpha and ranges between zero and 1, where the value of 1 indicates perfect reproducibility of person placements. The person reliability index is constructed using the measurement error and the observed variance associated with the person ability measures to calculate the ratio of "true" person ability variance to the observed variance. Also, the person separation index G_p is reported, which provides a measure of the sample standard deviation expressed in standard error units. The person separation index can be used to calculate the number of distinct ability strata H_p that can be reliably identified by the test scores according to the formula H_p.=$(4G_p+1)/3$.[191,195]

The ***targeting*** of the SIS domains is studied by examining the respective distribution of person abilities and item difficulties along the latent trait continuum. The distance between the mean person location and the mean item location, that is by definition set to zero, indicates domain targeting. Domains with mean person location values less than 0.5 logits below or beyond zero can be considered as well targeted. In addition, the percentage of persons with measures below the level of the lowest threshold, and of those with measures above the level of the highest threshold are calculated for each domain of the SIS to further evaluate domain targeting. For these cases the estimates of the quality of life in the concerning domain are imprecise, as they lay outside the scope of what can be measured by the items.

Differential item functioning (DIF) is examined to check for the invariance of calibrations, or item bias, across four dichotomous person factors: gender (male or female), age (< or >= 65 years), severity of stroke-related disability (Rankin score <= or > 2), and health care setting (inpatient or outpatient). Potential DIF is confirmed by analysis of the standardized residuals using a two-way analysis of variance (ANOVA) for each person factor. Two major types of DIF may be identified. Uniform DIF is shown by a difference in the residuals of the two groups, i.e. of male and female patients, corresponding to a mean effect of gender. In contrast, non-uniform DIF means that the residuals of persons are found to vary with group membership as well as ability levels, which corresponds to an interaction effect in the ANOVA results. Within each SIS domain a respective Bonferoni corrected type I error level of 0.01 will be used to identify significant DIF.[196,197]

To check whether the Rasch based item deletion conducted with results from a North American sample that resulted in the most current version of the SIS, the ***SIS 3.0***, can be confirmed in a German population of stroke patients, fit results of the Rasch analysis of the SIS 2.0 by Duncan et al. (2003) are compared descriptively to the results from the current study.

V.3. Results

The study population consists of a convenience sample of 196 eligible stroke patients from 16 study centers. Demographic and stroke-related characteristics of the study sample are presented in Table XIV. Using the standard scoring algorithm for the SIS, patients' domain scores have been calculated. According to the mean domain scores, the highest level of stroke-related QoL was observed in the domain Communication, while the lowest mean score is displayed at the domain Hand Function (Table XV). This is also reflected by the percentage of patients at the minimum and the maximum score levels respectively, where 29% of patients received the maximum score for Communication and 25% scored at the lowest possible score for Hand Function.

Table XVI summarizes basic person and item statistics for the eight SIS domains resulting from the first run of Rasch analyses prior to data purification. The unidimensionality of the single domains was explored first. Average infit mean square values across the domains range between .98 and 1.02. Table XVII shows the misfitting items and the according infit mean square statistics. Misfitting items have been identified in five out of the eight SIS domains displaying an infit larger than the preset criteria of 1.3. In the domains Emotion, Communication, Mobility, and Participation, respectively, one item shows misfit to the applied Rasch model ("enjoy things", "say name", "sit without loosing balance", "ability to show feelings"), whereas in the domain ADL/IADL three items display misfit, and thus, indicate deviation from unidimensionality ("control bladder", "handle money", "cut food with knife and fork").

The response category structure per domain was examined using the estimated threshold parameters for each item and checking their sequence. Altogether 25 items exhibit reversed thresholds. Only the SIS domain Strength shows in all items' category thresholds the expected pattern of increasing values. In contrast, 10 out of the 12 items within the domain ADL/IADL display category dysfunction. The number of items with disordered thresholds per SIS domain and the following data purification steps including collapsing of response categories and stepwise item deletion are summarized in Table XVIII.

In the first step of the data purification the ideal collapsing solution was selected for each domain resulting in ordered response categories for all items. For the SIS domains Memory/Thinking, Mobility, and Hand Function a four step response scale has been created. However, within the domain Memory/Thinking for the item "add and subtract numbers" only two threshold parameters were estimated due to the lack of responses in the category "1 – extremely difficult". For the SIS domains Emotion, Communication, ADL/IADL, and Participation a three step response scale resulted in ideal domain properties. However, in the domain Communication for the item “understand what is said” only one threshold parameter was estimated due to the lack of responses in the category “1 – extremely difficult”. After the collapsing of response categories, item fit was rechecked. Average infit mean square values across the domains range between .97 and 1.01. Three items that revealed discrepancy to the model prior to the purification procedures (“enjoy things”, “say name”, “cut food with knife and fork”) showed no misfit after the collapsing. Four items (“control bladder”, “handle money”, “sit without loosing balance”, “ability to show feelings”) still misfitted the model and were, therefore, deleted. During the stepwise item deletion within the domain ADL/IADL two further items (“control bowels”, “manage finances”) showed misfit and were deleted as well. Table XIX resumes person and item statistics for the eight SIS domains after the data purification procedure.

To examine measurement reliability, the person separation reliability values were considered. They range between .71 for the SIS domain Communication and .92 for the domain Mobility, as presented in Table XIX. For six out of the eight SIS domains the person separation reliability is .80 and above. Using SIS scores patients can be reliably distinguished into five separate strata in the domain Mobility, three separate strata in the domains Strength, Memory/Thinking, ADL/IADL, Mobility, Hand Function, and Participation, and in two separate strata in the domains Emotion and Communication. The data purification did not result in considerable changes of reliability.

The mean person location parameters for each domain following data purification (Table XIX) are examined to specify targeting. The domains Strength and ADL/IADL show mean person location values close to the mean item location of zero (.39 and .06, respectively). For Participation and Hand Function mean person locations lay more than 0.5 logits below the mean item location, i.e. overall the patients have lower levels of quality of life in these areas than covered by the items. In contrast, for the domains Communication and Memory/Thinking the mean person measures lay more than 3 logits above the mean item calibration, i.e. the patients' quality of life in these areas is at a higher level in this sample, than captured by the items. Table XX shows for each SIS domain the range of person measures, the range of the response category thresholds, as well as the percentage of persons falling below or beyond the logit range covered by the items' response categories as indicated by the respective thresholds. For the domains Memory/Thinking and Communication, 32% and 58% of the patients, respectively, have person measures falling beyond the measurement range of the items' response categories. For Hand Function and Participation, 29% and 21% of the patients displayed measures below the range of the category thresholds.

Differential item functioning (DIF) was examined across the four dichotomous person factors: gender, age, severity of stroke-related disability according to the Rankin Scale, and health care setting. Differential item functioning was not detected for any of the SIS items. Accordingly, all item locations can be regarded as invariant and unbiased with respect to the four examined factors.

Table XXI shows the comparison of item misfit in the current study of a German sample with the results of the Rasch analyses conducted by Duncan et al (2003)[148] with data from a North American population as basic reference. In the published analyses of Duncan et al (2003) three items showed misfit to the model in the eight SIS domains (infit >1.30) and were deleted to create the most current version of the SIS, the SIS 3.0. In the current study prior to data purification 7 items misfitted the model expectation. Item fit changed with the data purification procedure. During the data purification 6 items showed misfit and were deleted accordingly. One item, namely “ability to show feelings” from the domain Participation, could be identified, which met the misfit criterion in both studies. The two further items that misfitted in the reference study (“add and subtract numbers”, “get up from chair”) on the other hand accord with the model in the current study. In the current study five further items show misfit following data purifications, while they fitted in the study of Duncan et al (2003).

V.4. Discussion

The current study is the first psychometric evaluation of a stroke-specific QoL measure in Germany using a Rasch based approach. Applying Rasch analysis for instrument evaluation offers several advantages,[97,198] and can be regarded as a refinement of and an extension to traditional test theory. A Rasch based approach provides refined information on validity, e.g. when examining unidimensionality of the measured construct and the fit of the items using a reference that is external to the data. Furthermore, Rasch analysis also enables the evaluation of response category functioning and scale validity within a probabilistic framework. By yielding sample and test independent estimates of person and item parameters placed on the same single continuum, Rasch analysis makes possible a direct appraisal of test targeting, and also provides an index of reliability that is independent of sample distribution.

The results provide support for internal and construct validity, as well as reliability of the Stroke Impact Scale, and also point out issues for further improvement and adaptation of the SIS. The results of the current study add valuable information on the psychometric properties of the SIS when used in a German sample complementing the traditional test theory (TTT) based validation conducted by Petersen et al[188] and endorsing differential item functioning and response scale structure analyses, which have not been reported yet.

SIS data for the current study were collected from a heterogeneous sample including patients with different stroke etiologies, with different severity of disability, at different time points after stroke, and in different health care settings. According to the examination of differential item functioning, the SIS is applicable to different patient groups and yields valid and comparable measures unbiased by the person factors gender, age, disability severity and health care setting. Furthermore, using the SIS a high precision of measurement can be achieved as indicated by the high Person Separation Reliability indices that range above .80 for six of the SIS domains. Using SIS scores patients can be reliably distinguished into three to five separate strata in these domains. Reliability results are comparable to Cronbach's alpha values in previous studies[48,147,188] and are comparable to the Person Separation Reliability in the analysis of the original SIS 2.0.

Testing for model fit of the eight SIS domains confirmed unidimensionality according to the excellent mean infit statistics (.97 to 1.02). This result supports the findings from the TTT based study of the German SIS where high values of scale fit for all SIS domains were detected using confirmatory correlational MAP-analysis.[188] In the current study, the stepwise deletion of only six items out of 64 resolved model misfit and resulted in unidimensional domains. This item deletion has been conducted solely for statistical reasons to preserve the usefulness of the estimated parameters and the further results based on them, which is only given with ordered thresholds and fit to the model. Item reduction for the purpose of creating an improved version of an existent measure should be guided by further considerations and extensive study of the malfunctioning items. Further studies are required to understand the sources of item misfit before permanent item reduction can be conducted.

The examination of the response category functioning revealed a large number of 25 items with disordered thresholds. In the study of the original SIS 2.0 version by Duncan et al. (2003) response category functioning is not reported consistently and has not been accounted for with regard to the instruments sensitivity to change. Reversal of thresholds occurs, as patients' choice of response categories does not accord with the expectations from their estimated QoL level. This might happen because of ambiguity of the response categories in the understanding of the patients or because of the narrow range of the response options. Dysfunctioning of response categories is a major challenge to the interpretability of the SIS scores. In the current study threshold disorder has been accounted for post hoc by collapsing response categories in a way that preserved the precision of measurement and improved item fit. However, further studies are required to prove the usefulness of the reduced scales.

With regard to item fit the results of the original study of the SIS 2.0 by Duncan et al. (2003)[148] could not be replicated and thus, an item reduced version of the German SIS equivalent to the most current SIS 3.0 can not be created. Differing results in item fit might indicate cross-cultural differences in the meaning of the items, which should be object to further examination. Specific adaptation of the items in German speaking countries might be necessary. Rasch based examination of DIF across countries would be the method of choice in future studies to identify and account for cross-cultural differences in SIS items. The failure to replicate the fit results of the reference study might also be attributed to the use of the infit mean square statistic to detect and to compare misfit. As the infit value is an information-weighted statistic and the targeting of the test influences information, the differences in targeting in the different patient samples might have led to the observed discrepancy of the results in the two studies.

The SIS proved to be well targeted for a large proportion of moderately affected patients in the current sample, but also displayed floor and ceiling effects in some domains. As the SIS has been originally developed for patients with mild to moderate stroke, it is not exceptional that the QoL of patients at the lower levels of the measurement continuum was not covered in the domains Hand Function and Participation, i.e. for a high percentage of patients QoL in these areas is more severely affected than can be determined by the SIS. However, the QoL level of a large number of the patients located at the higher end of the measurement continuum also exceeds the coverage of the items, especially in the domains Memory/Thinking and Communication, but also in the domains Emotion and Mobility items seem to mark too low QoL levels with respect to the current sample. Thus, targeting of the SIS could be improved by adding both easier and harder items to avoid possible floor and ceiling effects and to appropriately represent the patients' levels of affection. However, stroke is a most heterogeneous condition as survivors state may range from vegetative state to no symptoms at all, thus no questionnaire is expected to be able to cover the full range of stroke-related problems. Rasch based analyses, as used in the current study, might become the method of choice in future to improve outcome measurement in the special population of stroke patients by facilitating the development of item banks and tailored testing.[101,199,200]

The current study used data from a comparatively small convenience sample of 196 stroke patients who were also selected by their capability to complete a self-report questionnaire. Thus, the sample does not include patients with low communication and cognitive abilities which is clearly reflected in the targeting results. However, the distribution of different stroke etiologies in the sample is comparable to that in the German population[201] and also different levels of stroke disability are well represented according to the Rankin Scale scores. A further limitation of the current study is that the composite Physical Function domain was not considered throughout the analyses and only the eight basic domains of the SIS were included.

In summary, results of the current study support the internal validity of the German version of the SIS with respect to the unidimensionality of the different QoL constructs. The SIS also seems to be well targeted for a large part of stroke patients whose QoL is moderately affected. The SIS represents a valid measure of QoL in patients' with different levels of disability severity, within an inpatient as well as an outpatient setting, across age groups and genders. However, the response categories currently used with the SIS should be object to further study and revision. With the differing results in the current study of the German SIS from the analysis of the original SIS 2.0 the issue of cross-cultural validation and adaptation emerges. The SIS is a well developed and according to its psychometric qualities sufficiently robust measure of stroke-specific QoL, suitable to capture important aspects of the consequences of stroke, and suitable for the application in international studies. However, the SIS needs to prove its cross-cultural validity in future.

Table XIII: Item examples and respective response categories for each domain of the SIS 2.0

SIS Domain	Items	Item example
Strength	4	In the past week, how would you rate the strength of your arm that was most affected by your stroke? A lot of strength (5) - Quite a bit of strength (4) - Some strength (3) - A little strength (2) - No strength at all (1)
Memory/ Thinking	8	In the past week, how difficult was it for you to remember things that people just told you? Not difficult at all (5) - A little difficult (4) - Somewhat difficult (3) - Very difficult (2) - Extremely difficult (1)
Emotion	9	In the past week, how often did you feel sad? None of the time (5) - A little of the time (4) - Some of the time (3) - Most of the time (2) - All of the time (1)
Communication	7	In the past week, how difficult was it to say the name of someone who was in front of you? Not difficult at all (5) - A little difficult (4) - Somewhat difficult (3) - Very difficult (2) - Extremely difficult (1)
ADL/ IADL	12	In the past 2 weeks, how difficult was it to dress the top part of your body? Not difficult at all (5) - A little difficult (4) - Somewhat difficult (3) - Very difficult (2) – Cannot do at all (1)
Mobility	10	In the past 2 weeks, how difficult was it to walk without losing your balance? Not difficult at all (5) - A little difficult (4) - Somewhat difficult (3) - Very difficult (2) – Cannot do at all (1)
Hand Function	5	In the past 2 weeks, how difficult was it to use your hand that was most affected by your stroke to tie a shoelace? Not difficult at all (5) - A little difficult (4) - Somewhat difficult (3) - Very difficult (2) – Cannot do at all (1)
Participation	9	In the past 4 weeks, how much of the time have you been limited in your social activities? None of the time (5) - A little of the time (4) - Some of the time (3) - Most of the time (2) - All of the time (1)

Table XIV: Demographic and stroke-related patient characteristics (N=196)

		Median **n**	**Range** **%**
Age		63 years	20 – 89
	<65 years	117	59.7%
	>= 65 years	79	40.3%
Gender	male	117	61.7%
	female	71	38.3%
Stroke etiology	hemorrhagic	27	13.8%
	ischemic	147	75.0%
	unspecified	22	11.2%
Chronicity		41 days	3 days – 32 years
	<=90 days	142	72.5
	>90 days	54	27.5%
Health care setting	inpatient	164	83.7%
	outpatient	32	16.3%
Rankin Scale disability grades	0 (no symptoms)	2	1.0%
	1 (not significant)	23	11.7%
	2 (slight)	58	29.6%
	3 (moderate)	57	29.1%
	4 (moderately severe)	38	19.4%
	5 (severe)	7	3.6%
	missing	11	5.6%
Full functioning prior stroke	yes	120	61.2%
	no	68	34.7%
	missing	8	4.1%

Table XV: SIS domain scores and percentage of patients scoring at the minimum or maximum score

SIS Domain	n*	Mean (SD)	min %	max %
Strength	194	52.5 (27.0)	3.1%	8.3%
Memory/ Thinking	195	82.7 (18.8)	-	22.1%
Emotion	196	72.2 (16.5)	-	4.1%
Communication	195	86.4 (17.0)	-	29.2%
ADL/ IADL	195	67.6 (24.9)	0.5%	6.7%
Mobility	195	66.1 (29.5)	2.1%	7.7%
Hand Function	194	47.8 (37.8)	25.3%	14.4%
Participation	186	50.6 (27.0)	1.6%	2.7%

* No score is generated for persons who answered less then 50% of the questions within one domain.

Table XVI: Summary of person and item statistics resulting from the Rasch analysis of the German SIS 2.0 prior to data purification

SIS Domain	Person				Item		
	Separation reliability	Separation index (Strata)	Mean	SE	SE	Logit range at middle category	Logit range of category thresholds
Strength	.82	2.31 (3)	.39	1.02	.13	-1.12 ; 1.25	-7.10 ; 5.57
Memory/Thinking	.83	2.22 (3)	3.30	.95	.14	-0.62 ; 0.48	-3.07 ; 4.26
Emotion	.71	1.56 (2)	1.22	.49	.09	-0.55 ; 0.36	-1.82 ; 2.40
Communication	.67	1.42 (2)	2.81	1.03	.13	-0.54 ; 0.38	-2.63 ; 2.56
ADL/IADL	.88	2.67 (4)	1.08	.52	1.02	-2.01 ; 1.78	-3.05 ; 3.69
Mobility	.92	3.33 (5)	1.36	.73	.14	-2.43 ; 1.90	-3.62 ; 5.42
Hand Function	.84	2.31 (3)	-.25	1.17	.13	-0.87 ; 0.49	-3.31 ; 3.75
Participation	.84	2.27 (3)	.04	.50	.09	-1.24 ; 0.91	-2.91 ; 1.68

Abbreviation: SE, standard error.

Table XVII: Misfitting items for the different domains of the German SIS 2.0 and according infit mean square statistics

SIS Domain	Item	Infit mean square
Emotion	enjoy things	1.33
Communication	say name	1.53
ADL/IADL	control bladder	1.55
	handle money	1.36
	cut food with knife and fork	1.32
Mobility	sit without loosing balance	1.67
Participation	ability to show feelings	2.17

Table XVIII: Data purification steps with the number of items with disordered thresholds per SIS domain, the selected response category collapsing options and the results of the stepwise item deletion

SIS Domain	Step 0 Number of items with disordered category thresholds	Step 1 Selected category collapsing option	Step 2 Stepwise item deletion Deleted items	 Infit mean square
Strength	-	-	-	
Memory/ Thinking	1	12334	-	
Emotion	1	11223	-	
Communication	4	12223	-	
ADL/ IADL	10	11223	control bladder	1.59
			control bowels	1.59
			handle money	1.60
			manage finances	1.62
Mobility	3	12234	sit without loosing balance	1.79
Hand Function	1	12234	-	
Participation	5	11223	ability to show feelings	1.95

Table XIX: Summary of person and item statistics resulting from the Rasch analysis of the German SIS 2.0 after data purification

SIS Domain	Person				Item		
	Separation reliability	Separation index (Strata)	Mean	SE	SE	Logit range at middle category	Logit range of category thresholds
Strength	.82	2.31 (3)	.39	1.02	.13	-1.12 ; 1.25	-7.10 ; 5.57
Memory/Thinking	.81	2.09 (3)	3.94	1.07	.18	-0.72 ; 0.50	-3.36 ; 5.11
Emotion	.72	1.60 (2)	1.08	.73	.15	-1.12 ; 0.78	-2.44 ; 3.24
Communication	.71	1.56 (2)	3.39	1.22	.21	-1.56 ; 2.35	-4.81 ; 3.43
ADL/IADL	.84	2.26 (3)	.06	.98	.17	-1.84 ; 2.37	-4.32 ; 4.39
Mobility	.92	3.34 (5)	.87	.87	.17	-1.90 ; 2.08	-6.67 ; 5.43
Hand Function	.84	2.28 (3)	-.69	1.28	.17	-1.10 ; 0.57	-4.63 ; 3.88
Participation	.80	1.98 (3)	-.97	.99	.17	-0.87 ; 1.22	-2.90 ; 2.34

Abbreviation: SE, standard error.

Table XX: Targeting after data purification indicated by the range of person measures, the range of the response category thresholds, the percentage of persons falling below or beyond the logit range covered by the items' response categories for each of the SIS domains

SIS Domain	Person measure range		Category threshold range		Number (%) of persons below/ beyond the category threshold range	
	Min	Max	Min	Max	Below	Beyond
Strength	-8.27	7.21	-7.10	5.57	5 (2.6%)	19 (9.7%)
Memory/ Thinking	-3.15	7.38	-3.36	5.11	-	62 (31.8%)
Emotion	-3.88	5.36	-2.44	3.24	4 (2.0%)	12 (6.1%)
Communication	-4.26	5.83	-4.81	3.43	-	104 (58.1%)
ADL/ IADL	-6.24	5.97	-4.32	4.39	24 (12.3%)	19 (9.7%)
Mobility	-8.03	7.18	-6.67	5.43	9 (4.6%)	17 (8.7%)
Hand Function	-6.26	5.95	-4.63	3.88	56 (28.9%)	38 (19.6%)
Participation	-5.14	5.03	-2.90	2.34	39 (20.5%)	19 (10.0%)

Table XXI: Item misfit in the current study and in the reference study by Duncan et al (2003)

SIS Domain	Item	Current study		Study by Duncan et al (2003)[148]
		Infit prior collapsing	Infit post collapsing	Infit
Memory/ Thinking	add subtract numbers			1.51
Emotion	enjoy things	1.33		
Communication	say name	1.53		
ADL/IADL	control bladder	1.55	1.59	
	handle money	1.36	1.60	
	cut food with knife and fork	1.32		
	control bowels		1.59	
	manage finances		1.62	
Mobility	sit without loosing balance	1.67	1.79	
	get up from chair			1.50
Participation	ability to show feelings	2.17	1.95	1.56

VI. DISCUSSION:

Towards a unified measurement approach in stroke

In the current doctoral thesis, two complementary principles towards the description of individuals' burden and disability have been introduced: the health status measurement and the classification approach. At the conjunction of these conceptual approaches two methodological procedures have been regarded: the linking method and the Rasch method. The application of both methods has been illustrated. It is argued that the connection of these two conceptual principles and both methodological procedures bears the potential to lead to a unified measurement approach in stroke, based upon a common and comprehensive understanding of patients' functioning and health and at the same time based upon individual, objective measurement.

The basic advantage of the health status measurement approach is the provision of operationalizations for a great variety of concepts. However, the high number of existing generic, domain- and disease-specific instruments, the frequently indefinite and unconnected underlying conceptualizations of the instruments, the focus on specific and narrow contents, and the unknown or even deficient psychometric qualities represent common drawbacks in stroke health status measurement.[21,32,42,43]

The classification approach, represented by the *International Classification of Functioning, Disability and Health (ICF)* of the WHO, provides the advantage of a largely accepted, unified, and universal conceptual framework attended by a comprehensive, multi-purpose taxonomy. However, the classification approach needs to be tailored and adopted to the specific application situations, e.g. to specific disease conditions or health care settings.[64,65,66,202,203]

An explicit connection between the health status measurement and the classification approach is established by the linking method. The linking method bears the potential to integrate the advantages of both approaches and to mitigate their disadvantages. It can connect a comprehensive common framework and language with the operationalization of stroke related functioning and health.

In this context, the linking method and the Rasch method complement each other. Both can work together in various application fields to enhance and facilitate the description and assessment of disability in stroke. First, both methods can be used for the selection of measures, second, for the development of new or modified instruments, and third, within the emerging new paradigm of measurement: item banking and adaptive testing.[98,101,199,200]

The examination of health status measures based on the linking method and using the ICF based quantitative indices and qualitative descriptions can be useful in a variety of ways for the selection of measures. The content related features of the measures can be compared with each other to ensure that the relevant contents are covered and the content structure is appropriate for the application objective. In practice, also the purpose of a planned data collection (e.g. the research hypothesis to be tested or the clinical intervention aims to be achieved) can be expressed using the category system of the ICF.[92,168] In this way, the purpose of the investigation on the one hand and assessment instruments on the other hand can be matched to each other one to one based on the ICF. Moreover, the results of the linking of the instruments can be compared with the ICF Core Set for Stroke. In this way, the comprehensiveness of one instrument or the effectiveness of a combination of instruments in describing the full scope of relevant problems in stroke patients' functioning can be evaluated. The comparison of the instruments' content with the ICF Core Set for Stroke can be used to indicate and compare their content validity.

While content is the first and most important concern in instrument selection, content related considerations have to be complemented by strict psychometric evaluation. Rasch based psychometric methods represent a refinement and extension of traditional test theory yielding advanced parameters independent of sample distribution with known precision and bias, which are comparable across patient populations, time points, and application situations.[100,101] Knowledge of instruments' psychometric features allows conclusions on their interpretability and applicability.[204] Thus, both methods are necessary counterparts in supporting the well founded choice of instruments, the sound planning of data collections, and the meaningful interpretation of their results.

The two methods, the linking method and the Rasch method, can be used to facilitate the development of new or modified measures. The comparison of instruments contents with the ICF Core Set for Stroke can be used to identify important areas of stroke patients' functioning and health, which are scarcely captured by existing measures, yet. This can serve to guide further instrument development. Additionally, the content diversity and density indices derived from the ICF based content examination may point to measures with a potential for refinements as to clarity or redundancy. They further might be helpful to identify measures that due to their content structure are readily suitable for translation into different languages and use in international trials. While several measures have already been developed, which rely on the ICF's basic biopsychosocial framework,[42,47,87] in instrument development, also the taxonomy of the classification might be useful, for example, at the stage of item generation.

Rasch analyses within the frame of probabilistic test theory are among the current standard psychometric methods of instrument development.[191,205] Results of Rasch based methods provide feedback on the structure and dimensionality of the items representing the measurement construct, on measurement invariance in different patient groups, on reliability, targeting and response scale functioning. Rasch analysis facilitates item selection according to model fit and according to the range of item difficulties. Using Rasch analysis the intended performance of an instrument can be tested pursuing the principle of a theory-practice dialogue. In consequence, it is possible to carry out specific and purposeful modifications of an instrument derived from the results of Rasch analyses.

While Rasch based methods provide the means for instrument development and modification based on quantitative metric properties and content structure, the ICF based linking method contributes qualitative and semi-quantitative content information. The qualitative ICF based linking method may give the direction of instrument development and modification, the quantitative Rasch based method on the other hand provides the means to arrive at objective health status measurement.

Ultimately, Rasch analyses and the linking method can interact to implement a new paradigm in stroke health status measurement: item banking and adaptive testing.[98,101,199,200] Item banks are collections of items organized according to their content and according to an underlying theory or model.[206] The WHO's biopsychosocial model, the ICF's category system, and the *ICF Core Sets for Stroke* can serve as organizational frame for creating an item bank. The linking method readily yields the necessary connection between the ICF and the items of assessment instruments. Items within an item bank are calibrated as to their difficulty level using Rasch analyses in a consistent and comparable way. Following their calibration, in practical measurement items can be flexibly chosen according to their content as well as according to their difficulty level to match the purposes of the assessment situation, and the level of affection in a specific individual at a given time, respectively. This is the essence of adaptive testing. In computer adaptive testing

(CAT) the process of this selection is conducted automatically during the assessment of an individual according to a predefined algorithm. In this way, the ICF and Rasch methods can complement each other as fundamental stepping-stones to implement comprehensive, yet practical ICF based measurement of health status.

The advantages of these methods are especially relevant in the field of stroke with its heterogeneous aftermaths, with its wide variety of severity, and with the highly dynamic changes in stroke patients' functioning in the course of time.[10] While no single rating scale, questionnaire or performance test can cover the full scope of stroke patients' functioning and health experience, still, these methods enable individually tailored, comprehensive, comparable, and efficient measurement.[101,199,200]

From the perspective of the classification approach the question ***what to measure*** has been posed. From the perspective of the health status measurement approach the question is ***how to measure***. Both questions can be answered by applying the methods illustrated within this doctoral thesis, the linking method and the Rasch method. However, the connection of the classification approach with the health status measurement approach entails advantages reaching beyond the applications presented here. Advances of the classification approach, like the development of the *ICF Core Sets for Stroke*, and advances in health status measurement, like the application of Rasch analyses can be concatenated by the linking method. Rasch analysis and the linking method representing qualitative and quantitative methods, may shed light on different facets of stroke measurement, which combined increase information value and lead to a complete picture of functioning and health.

From this concatenation of different conceptual and methodological approaches, unified and comparable, conceptually sound, high quality measurement of functioning can emerge. The integration of a common reference framework with the merits of objective measurement within the proceedings of item banking and adaptive testing can contribute to compass a common standard and agreement on what and how to measure. A unified measurement approach could thereby be achieved in stroke. Advanced measurement can serve to promote precise, comprehensive, and efficient knowledge of stroke disability at the individual and at population levels, to enable better decisions for treatment and action, in the long run improving stroke care and relieving the burden to the patients.

VII. SUMMARY

1. Background

Stroke is a frequently occurring condition and a common cause of death and disability. Many stroke survivors are facing long-term ***disability***. The consequences of stroke on patients' functioning are usually complex and heterogeneous. Precise knowledge of patients' stroke related disability is necessary in health services provision and research. Clinical stroke management, but also epidemiological and clinical research, depend on the careful detection of functioning problems, as well as resources, in patients with stroke.

Two conceptual approaches to describe patients' disability can be distinguished: the health status measurement and the classification approach. ***Health status measures***, like standardized performance tests, rating scales, and questionnaires are used to operationalize and to assess patients' burden of disease, functioning and health.

The ***classification approach*** towards the description of patients' health state is represented by the World Health Organization's ***International Classification of Functioning, Disability and Health (ICF)****.* The ICF provides a comprehensive conceptual framework and a unified standardized language to describe health and health related states, both at the individual, as well as at population levels. To enhance the applicability of the classification, *ICF Core Sets* for specific health conditions have been developed in an evidence based and consensus based process.[51,65] The ***ICF Core Sets for Stroke*** are selections of salient ICF categories out of the whole classification, which describe the spectrum of problems in stroke patients' functioning based on the universal language of the ICF. The *ICF Core Sets* *for Stroke* represent the practical implementation of the classification approach in clinical practice and research.

The two approaches to represent stroke related disability, the health status measurement and the classification approach, can be regarded as complementary principles. From the classification perspective, the ICF and the ICF Core Sets can serve as standards to define ***what*** to measure. From the perspective of health status measurement the question ***how*** to measure can be answered.

An explicit connection between the two approaches can be established by the so-called ***linking method***.[91,92] Thereby, using the ICF's category system the contents of measures can be mapped, explored and compared in a standardized, transparent and straightforward way. The linking method can be useful for various purposes. The application of the linking method along with the ICF Core Sets constitutes a new approach for examining health status measures' content validity.

However, beyond content validity, meaningful measurement essentially depends on the psychometric quality of the applied instruments. Techniques based on modern test theory, especially ***Rasch analysis***, are increasingly adopted to ensure instruments' psychometric properties.

2. Objectives

In the following, the doctoral thesis is subdivided into four parts. The first three parts present different studies performed to pursue the objectives named below. Each of the three studies contains a respective discussion section referring to the results of the study. The fourth part of the doctoral thesis refers to aim four, namely the discussion of the relationship between the methods presented in the previous three parts.

The current doctoral thesis aims

(1) to illustrate, how the connection between the health status measurement approach and the classification approach can be established by the application of the linking method,

(2) to demonstrate, how this approach can be used to select health status measures based on their content validity,

(3) to show, how the psychometric features of health status measures can be examined based on Rasch analyses, and

(4) to discuss the relationship between the demonstrated methods in the context of the connection of the health status measurement and the classification approach.

3. Applying the linking method:

Content comparison of patient-centered health status measures used in stroke based on the International Classification of Functioning, Disability and Health (ICF)

The first study, *"Content comparison of patient-centered health status measures used in stroke based on the International Classification of Functioning, Disability and Health (ICF)"* illustrates the application of the linking method in stroke measurement. The objective of this study was to examine and to compare the contents of patient-centred health status measures used in stroke.

The specific aims of the study included the identification of generic and condition-specific patient-centred health status measures applied in stroke patients, the examination of the contents of the single measures based on their linking to the ICF, and the comparison of the contents of generic and stroke-specific measures.

A systematic literature review was conducted to identify current generic and condition-specific patient-centred health status measures applied in stroke. The most frequently used instruments were selected. The contents of the selected measures were examined by linking the concepts within the instruments' items to the ICF.

Six generic and seven stroke-specific health status measures were selected. Within the selected instruments 979 concepts were identified. 200 different ICF categories were used to map these concepts. No single ICF category is contained in all instruments. Out of the total 200 different ICF categories used, 77 (40%) applied to only one of the 13 selected measures. Overall, the most frequently used category is b152 *emotional functions*' contained in 53 items from 10 instruments. Stroke-specific measures more often address mental functions, while the selected generic instruments more often include Environmental Factors.

The study provides an overview on current patient-centered health status measures in stroke and their covered contents. The results of the content comparison provide valuable information to facilitate and to account for the selection of appropriate instruments for specific purposes in clinical as well as research settings.

4. Selecting health status measures based on content validity: Comparison of stroke-specific health status measures with the *Comprehensive ICF Core Set for Stroke*

The aim of the second study is to demonstrate how the ICF as a fundamental reference can be used to select health status measures according to their content validity. The specific aims are (1) to examine the content validity of the selected stroke-specific health status measures by comparing them with the *Comprehensive*

ICF Core Set for Stroke, and (2) to discuss the selection of measures based on their coverage of the ICF Core Set.

Taking the results from the previous study, the seven stroke-specific patient-centered health status measures are involved in the current analyses and compared to the *Comprehensive ICF Core Set for Stroke*. Descriptive frequency analyses are conducted to indicate the instruments' bandwidth and specificity of content coverage regarding the ICF Core Set.

67 (52%) out of the 130 categories of the ICF Core Set are covered by at least one of the examined instruments. The single measures cover in total between 29% and 14% of the *Comprehensive ICF Core Set for Stroke*. Overall, 31 categories of the ICF Core Set are measured at the more specific 3rd and 4th levels by at least one of the seven instruments. All instruments cover Activity and Participation and Body Functions, but only two instruments address Environmental Factors. No categories of the ICF component Body Structures are contained in the examined instruments. In contrast the *Comprehensive ICF Core Set for Stroke* contains categories from all ICF components. The Quality of Life Instrument for Young Haemorrhagic Stroke Patients (HSQuale), the Stroke-specific Quality of Life Scale (SSQoL), and the Stroke Impact Scale[48,87] (SIS) represent the top three instruments according to bandwidth as well as specificity of content coverage.

The comparison of instruments against the *Comprehensive ICF Core Set for Stroke* can be used to characterize and compare measures' content validity. The examination and comparison of patient-centered health status measures' content validity accounting for the bandwidth and the specificity of content coverage can serve as a first step of selecting a measure. However, further features of the measures have to be considered. Especially, their psychometric properties have to be carefully examined to accomplish the well-founded choice of appropriate measures to assess stroke related health status.

5. Applying the Rasch method:

Evaluation of the Stroke Impact Scale using Rasch Analyses

The third study, the *"Evaluation of the Stroke Impact Scale using Rasch Analyses"* undertakes the psychometric evaluation of the Stroke Impact Scale 2.0 (SIS),[48,87,188] in a German sample adopting Rasch based techniques. The specific aims of the study were to examine (1) the unidimensionality of the SIS domains and item fit, (2) the structure of the response scales, (3) the targeting of the domains, (4) reliability, (5) differential item functioning (DIF) or item bias for relevant patient groups, and (6) to compare the fit results of this study with the Rasch analysis results of the SIS 2.0 in a North American sample which led to the creation of the most current version of the SIS, the SIS 3.0.

The Rasch analyses based on Master's Partial Credit model[191] has been carried out using data collected from stroke patients in Germany within an ongoing multicentric international study.[65] 196 stroke patients from 16 study centers participated in the study and completed the Stroke Impact Scale. Unidimensionality of the eight SIS domains was confirmed according to the mean infit statistics (.97 to 1.02). 7 items displayed model misfit. Response categories of 25 items showed threshold disordering. For the domains *Communication* and *Memory/Thinking* ceiling effects (>3 logits) became apparent. Reliability values lay above . 80 in six domains. No DIF was found as to age, gender, disability severity, and rehabilitation setting. Item fit results in the current study differed from those in the reference study of the SIS 2.0 in a North American sample.

The SIS is according to its psychometric qualities a sufficiently robust, valid and reliable measure of stroke-specific quality of life. It seems suitable to capture consequences of stroke in patients' with different levels of disability severity, within an inpatient as well as an outpatient setting, across age groups and genders. However, the response categories currently used with the SIS should be object to further study and revision. The fit results of the reference study of the SIS 2.0 could not be replicated and therefore, an item reduced version of the German SIS equivalent to the most current SIS 3.0 can not be created. Thus, the SIS needs to prove its cross-cultural validity in future.

6. Discussion:

Towards a unified measurement approach in stroke

Two complementary principles towards the description of disability have been introduced: the health status measurement and the classification approach. Connected to these conceptual approaches two methodological procedures have been regarded: the linking method and the Rasch method. The application of both methods has been illustrated.

The connection of the classification approach with the health status measurement approach entails advantages reaching beyond the applications presented here. Rasch analysis and the linking method, representing qualitative and quantitative methods, may shed light on different facets of stroke measurement, which combined increase information value and lead to a complete picture of functioning and health. Advances of the classification approach, like the development of the *ICF Core Sets for Stroke*, and advances in health status measurement, like the application of Rasch analyses can be concatenated by the linking method. From this concatenation of different conceptual and methodological approaches, unified and comparable, conceptually sound, high quality measurement of functioning can emerge.

The integration of a common reference framework with the merits of objective measurement within the proceedings of item banking and adaptive testing can contribute to compass a common standard and agreement on ***what*** and ***how*** to measure. A unified measurement approach could thereby be achieved in stroke. Advanced measurement can serve to promote precise, comprehensive, and efficient knowledge of stroke disability at the individual and at population levels, to enable better decisions for treatment and action, in the long run improving stroke care and relieving the burden to the patients.

VIII. ZUSAMMENFASSUNG

1. Hintergrund

Schlaganfall ist eine häufig auftretende Gesundheitsstörung und eine häufige Ursache von Tod und ***Behinderung***. Viele Überlebende eines Schlaganfalls müssen mit bleibenden Behinderungen rechnen. Die Folgen eines Schlaganfalls auf die Funktionsfähigkeit sind in der Regel komplex und vielseitig. In der Gesundheitsversorgung wie in der Forschung ist präzises Wissen um schlaganfallbezogene Behinderungen unerlässlich. Das klinische Management des Schlaganfalls, epidemiologische und klinische Forschung, sind angewiesen auf die sorgfältige Erfassung von Problemen und Ressourcen in der Funktionsfähigkeit der PatientInnen nach einem Schlaganfall.

Zwei konzeptionelle Ansätze zur Beschreibung von Behinderungen können unterschieden werden: der Ansatz der Messung des Gesundheitszustands (health status measurement) und der klassifizierende Ansatz. ***Messverfahren zur Erfassung des Gesundheitszustands***, wie etwa standardisierte Leistungstests, Ratingskalen und Fragebögen werden zur Operationalisierung und Messung der Krankheitsbelastung, der Funktionsfähigkeit und Gesundheit eingesetzt.

Der ***klassifizierende Ansatz*** zur Beschreibung des Gesundheitszustands ist durch die ***Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF)***[49] der Weltgesundheitsorganisation repräsentiert. Die ICF stellt einen umfassenden konzeptuellen Rahmen und eine einheitliche, standardisierte Sprache zur Beschreibung von Gesundheit und gesundheitsbezogenen Zuständen auf individueller aber auch auf Populationsebene zur Verfügung. Um die Anwendung der Klassifikation zu erleichtern wurden ICF Core Sets für bestimmte

Gesundheitsstörungen in einem evidenz-basierten und zugleich konsens-basierten Prozess entwickelt.[51,65] Die ***ICF Core Sets für Schlaganfall*** beinhalten eine Auswahl von relevanten ICF Kategorien aus der Gesamtklassifikation, die das Spektrum der Probleme in der Funktionsfähigkeit anhand der universellen Sprache der ICF beschreiben. Die ICF Core Sets für Schlaganfall stellen die praktische Umsetzung des klassifizierenden Ansatzes in der klinischen Praxis und Forschung dar.

Die zwei Ansätze zur Beschreibung von Behinderung bei SchlaganfallpatientInnen können als sich gegenseitig ergänzende Prinzipien betrachtet werden. Aus der Perspektive des klassifizierenden Ansatzes können die ICF und die ICF Core Sets als Standards dienen, um zu definieren, ***was*** erfasst werden sollte. Aus der Perspektive der Messung des Gesundheitszustands kann die Frage beantwortet werden, ***wie*** erfasst werden sollte.

Eine explizite Verbindung zwischen den beiden Ansätzen stellt die so genannte ***Linking Methode*** her.[91,92] Dabei können unter Anwendung des Kategoriensystems der ICF die Inhalte von Messverfahren in einer standardisierten, transparenten und direkten Weise abgebildet, untersucht und miteinander verglichen werden. Die Linking Methode kann für unterschiedliche Zwecke eingesetzt werden. Die Anwendung der Linking Methode zusammen mit den ICF Core Sets bildet einen neuen Zugangsweg für die Untersuchung der Inhaltsvalidität von Messverfahren zur Erfassung des Gesundheitszustands.

Neben der Inhaltsvalidität ist jedoch eine sinnvolle Messung im Wesentlichen von den psychometrischen Gütekriterien der verwendeten Instrumente abhängig. Techniken auf der Grundlage moderner Testtheorie, insbesondere ***Rasch Analysen***, werden zunehmend eingesetzt, um die psychometrischen Eigenschaften von Instrumenten zu überprüfen.

2. Ziele

Im Folgenden gliedert sich die Doktorarbeit in vier Teile. Die ersten drei Teile stellen unterschiedliche Untersuchungen, welche die unten genannten Zielsetzungen verfolgen dar. Jede dieser drei Untersuchungen beinhaltet einen eigenen Abschnitt, der sich auf die Diskussion der jeweiligen Ergebnisse bezieht. Der vierte Teil der Doktorarbeit befasst sich mit der vierten Zielsetzung, nämlich mit der Diskussion der Beziehungen zwischen den Methoden, die in den vorangegangenen Abschnitten vorgestellt wurden.

Die vorliegende Doktorarbeit hat zum Ziel,

(1) zu veranschaulichen, wie die Verbindung zwischen dem Ansatz der Messung des Gesundheitszustands und dem klassifizierenden Ansatz durch die Anwendung der Linking Methode hergestellt werden kann,

(2) zu demonstrieren, wie dieses Vorgehen genutzt werden kann, um Messverfahren aufgrund ihrer Inhaltsvalidität auszuwählen,

(3) zu zeigen, wie die psychometrischen Eigenschaften eines Messverfahrens mit Hilfe von Rasch Analysen untersucht werden können, und

(4) den Zusammenhang zwischen den vorgestellten Methoden im Kontext der Verbindung der beiden konzeptionellen Ansätze zu diskutieren.

3. Die Anwendung der Linking Methode:

Inhaltsvergleich patientenzentrierter Messverfahren zur Erfassung des Gesundheitszustands bei Schlaganfall anhand der Internationalen Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF)

Die erste Untersuchung, *“Inhaltsvergleich patientenzentrierter Messverfahren zur Erfassung des Gesundheitszustands bei Schlaganfall anhand der Internationalen Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit (ICF)“* zeigt die Anwendung der Linking Methode im Bereich Schlaganfall. Das Ziel dieser Studie war es, die Inhalte patientenzentrierter Messverfahren zur Erfassung des Gesundheitszustands nach einem Schlaganfall zu untersuchen und zu vergleichen.

Die spezifischen Ziele der Studie waren: die Identifikation bei SchlaganfallpatientInnen eingesetzter generischer und schlaganfallspezifischer, patientenzentrierter Verfahren zur Erfassung des Gesundheitszustands, die Untersuchung der Inhalte der einzelnen Messverfahren basierend auf ihrer Verknüpfung mit der ICF, sowie der Vergleich der Inhalte generischer und schlaganfallspezifischer Verfahren.

Ein systematischer Literaturreview wurde durchgeführt um generische und schlaganfallspezifische, patientenzentrierte Verfahren zur Erfassung des Gesundheitszustands zu identifizieren. Die am häufigsten bei SchlaganfallpatientInnen eingesetzten Instrumente wurden ausgewählt. Die Inhalte der ausgewählten Verfahren wurden untersucht indem die Konzepte innerhalb der Items der Instrumente mit der ICF verknüpft wurden.

Sechs generische und sieben schlaganfallspezifische Verfahren zur Erfassung des Gesundheitszustands wurden ausgewählt. In diesen Instrumenten wurden 979 Konzepte identifiziert. 200 verschiedene ICF Kategorien wurden verwendet, um die

Konzepte abzubilden. Keine einzige ICF Kategorie ist in allen Instrumenten enthalten. Von den insgesamt 200 verschiedenen genutzten ICF Kategorien waren 77 (40%) nur auf einen einzigen der 13 Instrumente anwendbar. Die insgesamt am häufigsten eingesetzte Kategorie war *b152 Emotionale Funktionen*, die in 53 Items aus 10 Instrumenten beinhaltet war. Schlaganfallspezifische Verfahren beinhalten häufiger Fragen nach mentale Funktionen, während die ausgewählten generischen Instrumente häufiger Umweltfaktoren einschließen.

Die Studie bietet einen Überblick über patientenzentrierte Messverfahren zur Erfassung des Gesundheitszustands nach einem Schlaganfall und deren Inhalte. Die Ergebnisse des Inhaltsvergleichs bieten wertvolle Informationen, um die Auswahl geeigneter Instrumente für die klinische Praxis und für die Forschung zu erleichtern und zu begründen.

4. Auswahl von Messverfahren zur Erfassung des Gesundheitszustands anhand ihrer Inhaltsvalidität:

Vergleich der schlaganfallspezifischen Messverfahren zur Erfassung des Gesundheitszustands mit dem Umfassenden ICF Core Set für Schlaganfall

Ziel der zweiten Untersuchung ist zu zeigen, wie die ICF als Basisreferenz für die Auswahl von Messverfahren zur Erfassung des Gesundheitszustands anhand ihrer Inhaltsvalidität genutzt werden kann. Die spezifischen Ziele der Studie waren: (1) die Inhaltsvalidität der schlaganfallspezifischen Messverfahren zur Erfassung des Gesundheitszustands zu untersuchen indem sie mit dem *Umfassenden ICF Core Set für Schlaganfall* verglichen werden, und (2) die Auswahl von Messinstrumenten anhand ihrer Abdeckung des ICF Core Sets zu diskutieren.

Unter Verwendung der Ergebnisse der vorangegangenen Untersuchung, werden die sieben schlaganfallspezifischen, patientenzentrierten Messverfahren zur Erfassung des Gesundheitszustands in die Analysen einbezogen und mit dem Umfassenden ICF Core Set für Schlaganfall verglichen. Deskriptive Häufigkeitsanalysen werden durchgeführt, um die Bandbreite und Spezifität der Inhaltsabdeckung bezüglich des ICF Core Sets abzubilden.

67 (52%) der 130 Kategorien des ICF Core Sets werden von mindestens einem der Instrumente abdeckt. Die einzelnen Messverfahren decken zwischen 29% und 14% des *Umfassenden ICF Core Sets für Schlaganfall* ab. Insgesamt werden 31 Kategorien des ICF Core Sets auf der spezifischeren dritten und vierten Ebene der Klassifikation durch mindestens einen der sieben Instrumente erfasst. Alle Instrumente decken Aktivitäten, Partizipation und Körperfunktionen ab, doch nur zwei Instrumente beinhalten Umweltfaktoren. Kategorien aus der ICF Komponente der Körperstrukturen sind in keinem der untersuchten Verfahren enthalten. Im Gegensatz dazu beinhaltet das *Umfassende ICF Core Set für Schlaganfall* Kategorien aus allen Komponenten der ICF. Der HSQuale - Quality of Life Instrument for Young Haemorrhagic Stroke Patients, der SSQoL - Stroke-specific Quality of Life Scale und der SIS - Stroke Impact Scale[48,87] sind die top drei Instrumente sowohl bezüglich Bandbreite als auch bezüglich Spezifität der Inhaltsabdeckung.

Der Vergleich der Instrumente mit dem *Umfassenden ICF Core Set für Schlaganfall* kann genutzt werden um die Inhaltsvalidität der Messverfahren zu charakterisieren und zu vergleichen. Die Untersuchung und Vergleich der Inhaltsvalidität von patientenzentrierten Messverfahren zur Erfassung des Gesundheitszustands unter Berücksichtigung der Bandbreite und der Spezifität der Inhaltsabdeckung kann als ein erster Schritt zur Auswahl eines Verfahrens dienen. Jedoch müssen weitere Eigenschaften der Instrumente berücksichtigt werden.

Insbesondere müssen die psychometrischen Gütekriterien sorgfältig untersucht werden, um eine wohlbegründete Wahl eines angemessenen Verfahrens zur Erfassung des schlaganfallbezogenen Gesundheitszustands zu treffen.

5. Die Anwendung der Rasch Methode: Evaluation der Stroke Impact Scale mittels Rasch Analysen

In der dritten Untersuchung, *“Evaluation der Stroke Impact Scale mittels Rasch Analysen”* wurden die psychometrischen Eigenschaften der Stroke Impact Scale 2.0 (SIS)[48,87,188] in einer deutschen Stichprobe unter Anwendung Rasch analytischer Techniken überprüft. Die spezifischen Ziele der Studie waren, (1) die Unidimensionalität der SIS Domänen und den Item Fit, (2) die Struktur der Antwortskalen, (3) Boden- und Deckeneffekte,(4) die Reliabilität, und (5) Differential Item Functioning (DIF) oder Item Bias in relevanten Patientengruppen zu untersuchen sowie (6) die Ergebnisse der vorliegenden Studie bezüglich Fit mit den Ergebnissen einer Rasch Analyse der SIS 2.0 in einer nord-amerikanischen Stichprobe zu vergleichen, die zur Erstellung der aktuellsten Version der SIS, der SIS 3.0 geführt hat.

Die Rasch Analyse basierend auf Masters Partial Credit Modell[191] wurde mit Daten aus einer deutschen Stichprobe durchgeführt, die innerhalb einer laufenden multizentrischen, internationalen Studie gesammelt worden sind.[65] 196 PatientInnen nach Schlaganfall aus 16 Studienzentren nahmen and der Studie teil und füllten die Stroke Impact Scale aus. Die Unidimensionalität der acht SIS Domänen wurde durch die mittleren Infit Werte (.97 to 1.02) bestätigt. 7 Items wiesen Abweichungen vom Rasch Modell auf. Bei 25 Items zeigten die Antwortkategorien ungeordnete Schwellenparameter. In den Domänen *Kommunikation* und *Gedächtnis/Denken* wurden Deckeneffekte (>3 logits) deutlich. Die Reliabilität lag in sechs Domänen bei über .80. Bezüglich Alter, Geschlecht, Schwere der Behinderung und Rehabilitationssetting wurde kein DIF gefunden. Die Ergebnisse bezüglich Item Fit unterschieden sich in der vorliegenden Studie von jenen der ursprünglichen Untersuchung der SIS 2.0 in einer nord-amerikanischen Stichprobe.

Die SIS ist ein bezüglich psychometrischer Eigenschaften ausreichend robustes, valides und reliables Messverfahren zur Erfassung der schlaganfall-spezifischen Lebensqualität. Sie scheint geeignet, Folgen eines Schlaganfalls, bei PatientInnen mit unterschiedlichen Schweregraden der Behinderung, in einem stationären als auch in einem ambulanten Setting, über Altersgruppen und Geschlecht hinweg einzufangen. Die Antwortkategorien der SIS sollten jedoch einer weiteren Untersuchung und Revision unterzogen werden. Die Ergebnisse der ursprünglichen Untersuchung der SIS 2.0 bezüglich Item Fit konnten nicht repliziert werden, sodass eine itemreduzierte Fassung der deutschen SIS, äquivalent zur aktuellsten SIS 3.0, nicht erstellt werden kann. Die SIS wird in der Zukunft ihre interkulturelle Validität noch unter Beweis stellen müssen.

6. Diskussion:

Auf dem Weg zu einem einheitlichen Messansatz bei Schlaganfall

Zwei komplementäre Grundsätze zur Beschreibung von Behinderung wurden vorgestellt: der Ansatz der Messung des Gesundheitszustands und der klassifizierende Ansatz. In Verbindung mit diesen konzeptionellen Ansätzen wurden zwei methodologische Vorgehensweisen berücksichtigt: die Linking Methode und die Rasch Methode. Die Anwendung beider Methoden wurde dargestellt.

Die Verbindung des klassifizierenden Ansatzes mit dem Ansatz der Messung des Gesundheitszustands, bringt Vorteile mit sich, die über die hier dargestellten Anwendungen hinausgehen. Rasch Analyse und die Linking Methode, die jeweils quantitative bzw. qualitative Methoden repräsentieren, können über unterschiedliche Facetten der Schlaganfallerfassung Aufschluss geben, können miteinander kombiniert den Nutzen der Informationen erhöhen und ein vollständiges Bild von Funktionsfähigkeit und Gesundheit zeichnen. Fortschritte beim klassifizierenden Ansatz, wie die Entwicklung der ICF Core Sets für Schlaganfall, und Fortschritte beim Ansatz der Messung des Gesundheitszustands, wie der Einsatz von Rasch Analysen, können durch die Linking Methode miteinander verbunden werden. Aus dieser Verbindung kann eine einheitliche, vergleichbare, konzeptionell schlüssige und qualitativ hochwertige Erfassung der Funktionsfähigkeit hervorgehen.

Die Integration eines gemeinsamen Bezugsrahmens mit den Vorteilen der objektiven Messung, wie es bei der Entwicklung von Item-Banken und adaptiven Tests vorgenommen werden kann, kann dazu beitragen, einen gemeinsamen Standard und eine Übereinstimmung darüber, ***was*** und ***wie*** erfasst werden sollte zu erreichen. Dadurch könnte ein einheitlicher Messansatz im Bereich Schlaganfall verwirklicht werden. Fortschrittliche Erfassungsmethoden können präzises, umfassendes und effizient nutzbares Wissen über schlaganfallbezogene Behinderungen auf individueller, wie auch auf Populationsebene erzeugen, bessere Entscheidungen über Maßnahmen und Behandlung ermöglichen und so auf lange Sicht zu einer Verbesserung der Schlaganfallversorgung und zu einer Linderung der Belastung der PatientInnen führen.

IX. LITERATURE

1 Mackay J, Mensah G. The Atlas of Heart Disease and Stroke. Geneva: World Health Organization; 2004.

2 American Heart Association. Heart disease and stroke statistics – 2003 update. Dallas, Texas: American Heart Association; 2002.

3 Thorvaldsen P, Asplund K, Kuulasmaa K, Rajakangas AM, Schroll M. Stroke incidence, case fatality, and mortality in the WHO MONICA project. Stroke 1995; 26: 361-367.

4 Kolominsky-Rabas PL, Heuschmann PU. Inzidenz, Ätiologie und Langzeitprognose des Schlaganfalls. Fortschr Neurol Psychiat 2002; 70: 657-662.

5 Muntner P, Garrett E, Klag MJ, Coresh J. Trends in stroke prevalence between 1973 and 1991 in the US population 25 to 74 years of age. Stroke 2002; 33: 1209-1213.

6 Kaste M, Fogelholm R, Rissanen A. Economic burden of stroke and the evaluation of new therapies. Public Health 1998; 112: 103-112.

7 Murray CJL, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet 1997; 349: 1436-1442.

8 Centers for Disease Control and Prevention. Prevalence of disabilities and associated health conditions among adults – United States, 1999. Morb Mortal Wkly Rep 2001; 50: 120-125.

9 World Health Organization. Stroke – 1989. Recommendations on stroke prevention, diagnosis, and therapy. Report of the WHO Task Force on stroke and other cerebrovascular disorders. Stroke 1989; 20: 1407-1431.

10 Jorgensen HS, Nakayama H, Raaschou HO, Vive-Larsen J, Stoier M, Olsen TS. Outcome and time course of recovery in stroke. Part II: Time course of

recovery. The Copenhagen Stroke Study. Arch Phys Med Rehabil 1995; 76: 406-412.

11 Jorgensen HS, Nakayama H, Raaschou HO, Vive-Larsen J, Stoier M, Olsen TS. Outcome and time course of recovery in stroke. Part I: Outcome. The Copenhagen Stroke Study. Arch Phys Med Rehabil 1995; 76: 399-405.

12 Hackett ML, Duncan JR, Anderson CS, Broad JB, Bonita R. Health-related quality of life among long-term survivors of stroke: results from the Auckland Stroke Study, 1991-1992. Stroke 2000; 31: 440-447.

13 Kase CS, Wolf PA, Kelly-Hayes M, Kannel WB, Beiser A, D'Agostino. Intellectual decline after stroke: the Framingham Study. Stroke 1998; 29: 805-812.

14 Srikanth VK, Thrift AG, Saling MM, Anderson JF, Dewey HM, Macdonell RA, Donnan GA. Increased risk of cognitive impairment 3 month after mild to moderate first-ever stroke. Stroke 2003; 34: 1136-1143.

15 Robinson RG. Neuropsychiatric consequences of stroke. Annu Rev Med 1997; 48: 217-229.

16 Whyte EM, Mulsant BH. Post-stroke depression: epidemiology, pathophysiology, and biological treatment. Biol Psychiatry 2002; 52: 253-264.

17 Stucki G, Ewert T, Cieza A. Value and application of the ICF in rehabilitation medicine. Disabil Rehabil 2002; 24: 932-938.

18 McHorney CA. Health status assessment methods for adults: past accomplishments and future challenges. Annu Rev Public Health 1999; 20: 309-335.

19 Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. Ann Intern Med 1993; 118: 622-629.

20 Bowling A. Measuring disease. A review of disease-specific quality of life measurement scales. Buckingham: Open University Press; 1995.

21 Buck D, Jacoby A, Massey A, Ford G. Evaluation of measures used to assess quality of life after stroke. Stroke 2000; 31: 2004-2010.

22 De Haan R, Aaronson N, Limburg M, Hewer RL, Van Crevel H. Measuring quality of life in stroke. Stroke 1993; 24: 320-327.

23 Golomb BA, Vickrey BG, Hays RD. A review of health-related quality-of-life measures in stroke. Pharmacoecomomics 2001; 19: 155-185.

24 Wade DT. Measurement in neurological rehabilitation. New York: Oxford University Press; 1992.

25 Brott T, Adams HP, Olinger CP, Marler JR, Barsan WG, Biller J, Spilker J, Holleran R, Eberle R, Hertzberg V. Measurements of acute cerebral infarction: a clinical examination scale. Stroke 1989; 20: 864-870.

26 Folstein MF, Folstein SE, McHugh PR. “Mini-Mental State”: a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975; 12: 189-198.

27 Ware JE, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30: 473-483.

28 EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. Health Policy 1990; 16: 199-208.

29 Nelson E, Wasson J, Kirk J, Keller A, Clark D, Dietrich A, Stewart A, Zubkoff M. Assessment of function in routine clinical practice: description of the COOP Chart method and preliminary findings. J Chronic Dis 1987; 40 (suppl 1): 55S-63S.

30 Geyh S, Kurt T, Brockow T, Cieza A, Ewert T, Omar Z, Resch KL. Identifying the concepts contained in outcome measures of clinical trials on stroke using the International Classification of Functioning, Disability and Health as a reference. J Rehabil Med 2004; 36 (Suppl 44): 56-62.

31 Roberts L, Counsell C. Assessment of clinical outcomes in acute stroke trials. Stroke 1998; 29: 986-991.

32 Duncan PW, Jorgensen HS, Wade DT. Outcome measures in acute stroke trials. A systematic review and some recommendations to improve practice. Stroke 2000; 31: 1429-1438.

33 Mahoney FI, Wood OH, Barthel DW. Rehabilitation of chronically ill patients: the influence of complications on the final goal. South Med J 1958; 51: 605-609.

34 Pound P, Gompertz P, Ebrahim S. A patient-centered study of the consequences of stroke. Clin Rehabil 1998; 12: 338-347.

35 Higginson IJ, Carr AJ. Measuring quality of life: Using quality of life measures in the clinical setting. BMJ 2001; 322: 1297-1300.

36 Mayo NE, Wood-Dauphinee S, Cote R, Durcan L, Carlton J. Activity, participation, and quality of life 6 months poststroke. Arch Phys Med Rehabil 2002; 83: 1035-42.

37 Kauhanen ML, Korpelainen JT, Hiltunen P, Nieminen P, Sotaniemi KA, Myllyla VV. Domains and determinants of quality of life after stroke caused by brain infarction. Arch Phys Med Rehabil 2000; 81: 1541-1546.

38 Hopman WM, Verner J. Quality of life during and after inpatient stroke rehabilitation. Stroke 2003; 34: 801-805.

39 Suenkeler IH, Nowak M, Misselwitz B, Kugler C, Schreiber W, Oertel WH, Back T. Timecourse of health-related quality of life as determined 3, 6 and 12 months

after stroke. Relationship to neurological deficit, disability and depression. J Neurol 2002; 249: 1160-1167.

40 Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. BMJ 2002; 324: 1417.

41 Sanders C, Egger M, Donovan J, Tallon D, Frankel S. Reporting on quality of life in randomised controlled trials: bibliographic study. BMJ 1998; 317: 1191-1194.

42 Doyle PJ. Measuring health outcomes in stroke survivors. Arch Phys Med Rehabil 2002; 83(12 Suppl 2): S39-43.

43 Williams LS. Health-related quality of life outcomes in stroke. Neuroepidemiology 1998; 17: 116-120.

44 Riemsma RP, Forbes CA, Glanville JM, Eastwood AJ, Kleijnen J. General health status measures for people with cognitive impairment: learning disability and acquired brain injury. Health Technol Assess 2001; 5: 1-100.

45 Williams LS, Weinberger M, Harris LE, Clark DO, Biller J. Development of a stroke-specific quality of life scale. Stroke 1999; 30: 1362-1369.

46 Buck D, Jacoby A, Massey A, Steen N, Sharma A, Ford GA. Development and validation of NEWSQOL, the Newcastle Stroke-specific Quality of Life Measure. Cerebrovasc Dis 2004; 17: 143-152.

47 Doyle PJ, McNeil MR, Mikolic JM, Prieto L, Hula WD, Lustig AP, Ross K, Wambaugh JL, Gonzalez-Rothi LJ, Elman RJ. The Burden of Stroke Scale (BOSS) provides valid and reliable score estimates of functioning and well-being in stroke survivors with and without communication disorders. J Clin Epidemiol 2004; 57: 997-1007.

48 Duncan PW, Wallace D, Lai SM, Johnson D, Embretson S, Laster LJ. The stroke impact scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. Stroke 1999; 30: 2131-2140.

49 World Health Organization. International Classification of Functioning, Disability and Health: ICF. Geneva: World Health Organization; 2001.

50 World Health Organization. International statistical classification of diseases and related health problems, 10th revision. Geneva: World Health Organization; 1992.

51 Stucki G, Cieza A, Ewert T, Kostanjsek N, Chatterji S, Ustun TB. Application of the International Classification of Functioning, Disability and Health (ICF) in clinical practice. Disabil Rehabil 2002; 24: 281-282.

52 World Health Organization. Towards a Common Language for Functioning, Disability and Health, ICF. Geneva: World Health Organization; 2002.

53 Steiner WA, Ryser L, Huber E, Uebelhart D, Aeschlimann A, Stucki G. Use of the ICF model as a clinical problem-solving tool in physical therapy and rehabilitation medicine. Phys Ther 2002; 82: 1098-1107.

54 Bilbao A, Kennedy C, Chatterji S, Ustun B, Barquero JL, Barth JT. The ICF: Applications of the WHO model of functioning, disability and health to brain injury rehabilitation. NeuroRehabilitation 2003; 18: 239-250.

55 Allan CM, Campbell WN, Guptill CA, Stephenson FF, Campbell KE. A conceptual model for interprofessional education: The international classification of functioning, disability and health (ICF). Journal of Interprofessional Care 2006; 20: 235-245.

56 Mau W, Gulich M, Gutenbrunner C, Lampe B, Morfeld M, Schwarzkopf SR, Smolenski UC. Lernziele im Querschnittsbereich Rehabilitation, Physikalische

Medizin und Naturheilverfahren nach der 9. Revision der Approbationsordnung für Ärzte--Gemeinsame Empfehlung der Deutschen Gesellschaft für Rehabilitationswissenschaften und der Deutschen Gesellschaft für Physikalische Medizin und Rehabilitation. Die Rehabilitation 2004, 43, 337-347.

57 Lollar DJ. Public health and disability: emerging opportunities. Public Health Rep 2002; 117: 131-136.

58 Bruyere SM, Van Looy SA, Peterson DB. The International Classification of Functioning, Disability and Health: Contemporary Literature Overview. Rehabilitation Psychology 2005; 50: 113–121.

59 Sato H. A survey on the use of ICF in policy. RIVM WHO-FIC Newsletter 2005; 3(1): 2.

60 Ewert T, Geyh S, Grill E, Cieza A, Zaisserer S, Stucki G. Die Anwendung der ICF in der Neurorehabilitation anhand des ICF- Modellblattes und der ICF Core Sets. Neurologie und Rehabilitation 2005; 11: 179-188.

61 W. Fries, D. Dustmann, S. Fischer, N. Lojewski, K. Ortner, C. Petersen, C. Pott, M. Rehbein, I. Scholler. Projektarbeit: Therapeutische Strategien zur Umsetzung von ICF und SGB IX in der ambulanten wohnortnahen neurologischen Rehabilitation zur Verbesserung der Teilhabe am Leben in der Gesellschaft. Neurologie und Rehabilitation 2005; 11: 218-226.

62 Rentsch HP, Bucher P, Dommen Nyffeler I, Wolf C, Hefti H, Fluri E, Wenger U, Walti C, Boyer I. The Implementation of the International Classification of Functioning, Disability and Health (ICF) in Daily Practice of Neurorehabilitation: an Interdisciplinary Project at the Kantonsspital of Lucerne, Switzerland. Disabil Rehabil 2003; 25: 411-421.

63 Rentsch HP, Bucher P: ICF in der Rehabilitation. Die praktische Anwendung der internationalen Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit im Rehabilitationsalltag. Idstein: Schulz-Kirchner Verlag; 2005.

64 Stucki G, Grimby G. Applying the ICF in medicine. J Rehabil Med 2004; (44 Suppl): 5-6.

65 Cieza A, Ewert T, Ustun TB, Chatterji S, Kostanjsek N, Stucki G. Development of ICF Core Sets for patients with chronic conditions. J Rehabil Med 2004; (44 Suppl): 9-11.

66 Üstün B, Chatterji S, Kostanjsek N. Comments from WHO for the Journal of Rehabilitation Medicine Special Supplement on ICF Core Sets. J Rehabil Med 2004; (44 Suppl): 7-8.

67 Brach M, Cieza A, Stucki G, Füssl M, Cole A, Ellerin BE, Fialka-Moser V, Kostanjsek N, Melvin J. ICF Core Sets for breast cancer. J Rehabil Med 2004; (44 Suppl): 121-127.

68 Cieza A, Stucki A, Geyh S, Berteanu M, Quittan M, Simon A, Kostanjsek N, Stucki G, Walsh N. ICF Core Sets for chronic ischemic heart disease. J Rehabil Med 2004; (44 Suppl): 94-99.

69 Cieza A, Stucki G, Weigl M, Kullmann L, Stoll T, Kamen L, Kostanjsek N, Walsh N. ICF Core Sets for chronic widespread pain. J Rehabil Med 2004; (44 Suppl): 63-68.

70 Cieza A, Chatterji S, Andersen C, Cantista P, Herceg M, Melvin J, Stucki G, de Bie R. ICF Core Sets for depression. J Rehabil Med 2004; (44 Suppl): 128-134.

71 Ruof J, Cieza A, Wolff B, Angst F, Ergeletzis D, Omar Z, Kostanjsek N, Stucki G. ICF Core Sets for diabetes mellitus J Rehabil Med 2004; (44 Suppl): 100-6.

72 Cieza A, Stucki G, Weigl M, Disler P, Jäckel W, van der Linden S, Kostanjsek N, de Bie R. ICF Core Sets for low back pain. J Rehabil Med 2004; (44 Suppl): 69-74.

73 Stucki A, Daansen P, Fuessl M, Cieza A, Huber E, Atkinson R, Kostanjsek N, Stucki G, Ruof J. ICF Core Sets for obesity. J Rehabil Med 2004; (44 Suppl): 107-113.

74 Stucki A, Stoll T, Cieza A, Weigl M, Giardini A, Wever D, Kostanjsek N, Stucki G. ICF Core Sets for obstructive pulmonary diseases. J Rehabil Med 2004; (44 Suppl): 114-120.

75 Dreinhöfer K, Stucki G, Ewert T, Huber E, Ebenbichler G, Gutenbrunner Ch, Kostanjsek N, Cieza A. ICF Core Sets for osteoarthritis. J Rehabil Med 2004; (44 Suppl): 75-80.

76 Cieza A, Schwarzkopf SR, Sigl T, Stucki G, Melvin J, Stoll T, Woolf A, Kostanjsek N, Walsh N. ICF Core Sets for osteoporosis. J Rehabil Med 2004; (44 Suppl): 81-86.

77 Stucki G, Cieza A, Geyh S, Battistella L, Lloyd J, Symmons S , Kostanjsek N, Schouten J. ICF Core Sets for rheumatoid arthritis. J Rehabil Med 2004; (44 Suppl): 87-93.

78 Geyh S, Cieza A, Schouten J, Dickson H, Frommelt P, Omar Z, Ring H, Kostanjsek N, Stucki G. *ICF Core Sets for Stroke*. J Rehabil Med 2004; (44 Suppl): 135-141.

79 Weigl M, Cieza A, Andersen C, Kollerits B, Amann E, Stucki G. Identification of relevant ICF categories in patients with chronic health conditions: A Delphi exercise. J Rehabil Med 2004; (44 Suppl): 12-21.

80 Brockow T, Cieza A, Kuhlow H, Sigl T, Franke T, Harder M, Stucki G. Identifying the concepts contained in outcome measures of clinical trials on musculoskeletal disorders and chronic wide spread pain using the International Classification of Functioning, Disability and Health as a reference. J Rehabil Med 2004; (44 Suppl): 30-36.

81 Wolff B, Cieza A, Parentin A, Rauch A, Sigl T, Brockow T, Stucki A. Identifying the concepts contained in outcome measures of clinical trials on four internal disorders using the International Classification of Functioning, Disability and Health as a reference. J Rehabil Med 2004; (44 Suppl): 37-42.

82 Brockow T, Duddeck K, Geyh S, Schwarzkopf SR, Weigl M, Franke T, Brach M. Identifying the concepts contained in outcome measures of clinical breast cancer trials Using the International Classification of Functioning, Disability and Health as a reference. J Rehab Med 2004; (44 Suppl): 43-48.

83 Brockow T, Wohlfahrt K, Hillert A, Geyh S, Weigl M, Franke T, Resch KL, Cieza A. Identifying the concepts contained in the outcome measures of trials on depressive disorders using the International Classification of Functioning, Disability and Health as a reference. J Rehab Med 2004; (44 Suppl): 49-55.

84 World Health Organization. ICF Checklist Version 2.1a, Clinical Form for International Classification of Functioning, Disability and Health: ICF. Geneva: World Health Organization; 2001.

85 Ewert T, Fuessl M, Cieza A, Andersen C, Chatterji S, Kostanjsek N, Stucki G. Identification of the most common patient problems in patients with chronic conditions using the ICF checklist. J Rehab Med 2004; (44 Suppl): 22-29.

86 World Health Organization. International Classification of Impairments, Disabilities and Handicaps. Geneva: World Health Organization; 1980.

87 Duncan PW, Wallace D, Studenski S, Lai SM, Johnson D. Conceptualization of a new stroke-specific outcome measure: the Stroke Impact Scale. Top Stroke Rehabil 2001; 8: 19-33.

88 Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J. Issues for selection of outcome measures in stroke rehabilitation: ICF Body Functions. Disabil Rehabil 2005; 27: 191-207.

89 Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: ICF Activity. Disabil Rehabil 2005; 27: 315-340.

90 Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. Issues for selection of outcome measures in stroke rehabilitation: ICF Participation. Disabil Rehabil 2005; 27: 507-528.

91 Cieza A, Brockow T, Ewert T, Amman E, Kollerits B, Chatterji S, Ustun TB, Stucki G. Linking health status measurements to the international classification of functioning, disability and health. J Rehabil Med 2002; 34: 205-210.

92 Cieza A, Geyh S, Chatterji S, Konstanjsek N, Üstün B, Stucki G. ICF linking rules: An update based on lessons learned. J Rehabil Med 2005; 37: 212-218.

93 Stucki G, Cieza A. The International Classification of Functioning, Disability and Health (ICF) Core Sets for rheumatoid arthritis: a way to specify functioning. Ann Rheum Dis 2004; 63 (Suppl 2): 40-45.

94 Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 2002; 11: 193-205.

95 Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess 1998; 2: 1-74.

96 Rasch G. Probabilistic models for some intelligence and attainment tests. Expanded edition. Chicago: University of Chicago Press; 1980.

97 McHorney CA, Monahan PO. Postscript: Applications of Rasch analysis in health care. Med Care 2004; 42 (1 Suppl): I73-I78.

98 Jette AM, Haley SM. Contemporary measurement techniques for rehabilitation outcomes assessment. J Rehabil Med 2005; 37: 339-345.

99 Cella D, Chang CH. A discussion of item response theory and its applications in health status assessment. Med Care 2000; 38 (9 Suppl): II66-II72.

100 Bond TG, Fox CM. Applying the Rasch model: Fundamental measurement in the human sciences. Mahwah, NJ: Lawrence Erlbaum Associates; 2001.

101 Andrich D. Rasch Models for Measurement. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-068. Newbury Park, CA: Sage; 1988.

102 Fujiwara T, Liu M, Tsuji T, Sonoda S, Mizuno K, Akaboshi K, Hase K, Masakado Y, Chino N. Development of a new measure to assess trunk impairment after stroke (trunk impairment scale): its psychometric properties. Am J Phys Med Rehabil 2004; 83: 681-688.

103 Duncan PW, Lai SM, Bode RK, Perera S, DeRosa J. Stroke Impact Scale-16: A brief assessment of physical function. Neurology 2003; 60: 291-296.

104 Engberg A, Garde B, Kreiner S. Rasch analysis in the development of a rating scale for assessment of mobility after stroke. Acta Neurol Scand 1995; 91: 118-127.

105 Hamilton BB, Granger CV, Sherwin FS, Zielezny M, Tashman JS. A uniform national data system for medical rehabilitation. In: Fuhrer MJ, editor.

Rehabilitation outcomes: analysis and measurement. Baltimore: Paul H. Brookes; 1987: 137-147.

106 Granger CV, Hamilton BB, Linacre JM, Heinemann AW, Wright BD. Performance profiles of the functional independence measure. Am J Phys Med Rehabil 1993; 72: 84-89.

107 Hartingsveld F, Lucas C, Kwakkel G, Lindeboom R. Improved interpretation of stroke trial results using empirical Barthel item weights. Stroke 2006; 37: 162-166.

108 Heinemann AW, Harvey RL, McGuire JR, Ingberman D, Lovell L, Semik P, Roth EJ. Measurement properties of the NIH Stroke Scale during acute rehabilitation. Stroke 1997; 28: 1174-1180.

109 Lundgren-Nilsson A, Grimby G, Ring H, Tesio L, Lawton G, Slade A, Penta M, Tripolski M, Biering-Sorensen F, Carter J, Marincek C, Phillips S, Simone A, Tennant A. Cross-cultural validity of functional independence measure items in stroke: a study using Rasch analysis. J Rehabil Med 2005; 37: 23-31.

110 Tennant A, Penta M, Tesio L, Grimby G, Thonnard JL, Slade A, Lawton G, Simone A, Carter J, Lundgren-Nilsson A, Tripolski M, Ring H, Biering-Sorensen F, Marincek C, Burger H, Phillips S. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. Med Care 2004; 42 (1 Suppl): I37-I48.

111 Kucukdeveci AA, Yavuzer G, Elhan AH, Sonel B, Tennant A. Adaptation of the Functional Independence Measure for use in Turkey. Clin Rehabil 2001; 15: 311-319.

112 Grimby G, Gudjonsson G, Rodhe M, Sunnerhagen KS, Sundh V, Ostensson ML. The functional independence measure in Sweden: experience for outcome measurement in rehabilitation medicine. Scand J Rehabil Med 1996; 28: 51-62.

113 Tsuji T, Sonoda S, Domen K, Saitoh E, Liu M, Chino N. ADL structure for stroke patients in Japan based on the functional independence measure. Am J Phys Med Rehabil 1995; 74: 432-438.

114 Nilsson AL, Sunnerhagen KS, Grimby G. Scoring alternatives for FIM in neurological disorders applying Rasch analysis. Acta Neurol Scand. 2005, 111: 264-273.

115 Microsoft Corporation. Microsoft Access 2000. Redmond, WA: Microsoft Corporation; 1992-1999.

116 Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. 1960; 20: 37-46.

117 Vierkant RA. A SAS macro for calculating bootstrapped confidence intervals about a Kappa coefficient. SAS Users Group International Online Proceedings 2000. Retrieved July 23, 2004, from <http://www2.sas.com/ proceedings/ sugi22/ STATS/ PAPER295.PDF>.

118 SAS Institute Inc. The SAS System for Windows, Version 8.2. Cary, NC: SAS Institute Inc; 2001.

119 Wood-Dauphinee SL, Opzoomer MA, Williams JI, Marchand B, Spitzer WO. Assessment of global function: The Reintegration to Normal Living Index. Arch Phys Med Rehabil 1988; 69: 583-590.

120 Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. Med Care 1981; 19: 787-805.

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

122 Hunt SM, McEwen J, McKenna SP. Measuring health status: a new tool for clinicians and epidemiologists. J R Coll Gen Pract 1985; 35: 185-188.

123 Hilari K, Byng S. Measuring quality of life in people with aphasia: the Stroke-specific Quality of Life Scale. Int J Lang Commun Disord 2001; 36 (Suppl): 86-91.

124 Ferrans C, Powers M. Quality of Life Index: Development and psychometric properties. Advances in Nursing Science 1985; 8: 15-24.

125 Van Straten A, de Haan RJ, Limburg M, Schuling J, Bossuyt PM, van den Bos GA. A stroke-adapted 30-item version of the Sickness Impact Profile to assess quality of life (SA-SIP30). Stroke 1997; 28: 2155-2161.

126 Hamedani AG, Wells CK, Brass LM, Kernan WN, Viscoli CM, Maraire JN, Awad IA, Horwitz RI. A quality-of-life instrument for young hemorrhagic stroke patients. Stroke 2001; 32: 687-695.

127 Stamm TA, Cieza A, Machold K, Smolen JS, Stucki G. An exploration of the link of conceptual occupational therapy models to the International Classification of Functioning, Disability and Health (ICF). Arthritis Rheum. In press.

128 Stucki G, Cieza A, Ewert T, Kostanjsek N, Chatterji S, Üstün B. Application of the International Classification of Functioning, Disability and Health (ICF) in clinical practice. Disabil Rehabil 2002; 24: 281-282.

129 Dorman P, Slattery J, Farrell B, Dennis M, Sandercock P. Qualitative comparison of the reliability of health status assessments with the EuroQol and SF-36 questionnaires after stroke. United Kingdom Collaborators in the International Stroke Trial. Stroke 1998; 29: 63-68.

130 Anderson C, Laubscher S, Burns R. Validation of the Short Form 36 (SF-36) health survey questionnaire among stroke patients. Stroke 1996; 27: 1812-1816.

131 Hagen S, Bugge C, Alexander H. Psychometric properties of the SF-36 in the early post-stroke phase. J Adv Nurs 2003; 44: 461-468.

132 Hobart JC, Williams LS, Moran K, Thompson AJ. Quality of life measurement after stroke: uses and abuses of the SF-36. Stroke 2002; 33: 1348-1356.

133 Clarke PJ, Black SE, Badley EM, Lawrence JM, Williams JI. Handicap in stroke survivors. Disabil Rehabil 1999; 21: 116-123.

134 Tooth LR, McKenna KT, Smith M, O'Rourke PK. Reliability of scores between stroke patients and significant others on the Reintegration to Normal Living (RNL) Index. Disabil Rehabil 2003; 25: 433-440.

135 Sneeuw KC, Aaronson NK, de Haan RJ, Limburg M. Assessing quality of life after stroke. The value and limitations of proxy ratings. Stroke 1997; 28: 1541-1549.

136 De Haan R, Horn J, Limburg M, Van Der Meulen J, Bossuyt P. A comparison of five stroke scales with measures of disability, handicap, and quality of life. Stroke 1993; 24: 1178-1181.

137 Dorman PJ, Waddell F, Slattery J, Dennis M, Sandercock P. Is the EuroQol a valid measure of health-related quality of life after stroke? Stroke 1997; 28: 1876-1882.

138 Pickard AS, Johnson JA, Feeny DH, Shuaib A, Carriere KC, Nasser AM. Agreement between Patient and Proxy Assessments of Health-Related Quality of Life after Stroke using the EQ-5D and Health Utilities Index. Stroke 2004; 35: 607-612.

139 Dorman PJ, Waddell F, Slattery J, Dennis M, Sandercock P. Are Proxy Assessments of Health Status After Stroke With the EuroQol Questionnaire Feasible, Accurate, and Unbiased? Stroke 1997; 28: 1883-1887.

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

141 Jenkinson C, Mant J, Carter J, Wade D, Winner S. The London handicap scale: a re-evaluation of its validity using standard scoring and simple summation. J Neurol Neurosurg Psychiatry 2000; 68: 365–367.

142 Gompertz P, Pound P, Ebrahim S. The reliability of stroke outcome measures. Clin Rehabil 1993; 7: 290-296.

143 Ebrahim S, Barer D, Nouri F. Use of the Nottingham Health Profile with patients after a stroke. J Epidemiol Community Health 1986; 40: 166-169.

144 Visser MC, Koudstaal PJ, Erdman RA, Deckers JW, Passchier J, van Gijn J, Grobbee DE. Measuring quality of life in patients with myocardial infarction or stroke: a feasibility study of four questionnaires in The Netherlands. J Epidemiol Community Health 1995; 49: 513-517.

145 Duncan PW, Lai SM, Tyler D, Perera S, Reker DM, Studenski S. Evaluation of proxy responses to the Stroke Impact Scale. Stroke 2002; 33: 2593-2599.

146 Duncan PW, Reker DM, Horner RD, Samsa GP, Hoenig H, LaClair BJ, Dudley TK. Performance of a mail-administered version of a stroke-specific outcome measure, the Stroke Impact Scale. Clin Rehabil 2002; 16: 493-505.

147 Edwards B, O'Connell B. Internal consistency and validity of the Stroke Impact Scale 2.0 (SIS 2.0) and SIS-16 in an Australian sample. Qual Life Res 2003; 12: 1127-35.

148 Duncan PW, Bode RK, Min Lai S, Perera S; Glycine Antagonist in Neuroprotection Americans Investigators. Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. Arch Phys Med Rehabil 2003; 84: 950-63.

149 Hilari K, Byng S, Lamping DL, Smith SC. Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): evaluation of acceptability, reliability, and validity. Stroke 2003; 34: 1944-1950.

150 Robinson-Smith G, Johnston MV, Allen J. Self-care self-efficacy, quality of life, and depression after stroke. Arch Phys Med Rehabil 2000; 81: 460-464.

151 King RB. Quality of Life after Stroke. Stoke 1996; 27: 1467-1472.

152 Van de Port IG, Ketelaar M, Schepers VP, Van den Bos GA, Lindeman E. Monitoring the functional health status of stroke patients: the value of the Stroke-Adapted Sickness Impact Profile-30. Disabil Rehabil 2004; 26: 635-640.

153 Doyle PJ, McNeil MR, Hula WD, Mikolic JM.The Burden of Stroke Scale (BOSS): Validating patient-reported communication difficulty and associated psychological distress in stroke survivors. Aphasiology 2003; 17: 291-304.

154 Rochette A, Desrosiers J, Noreau L. Association between personal and environmental factors and the occurrence of handicap situations following a stroke. Disabil Rehabil 2001; 23: 559-569.

155 Evans RL, Connis RT, Bishop DS, Hendricks RD, Haselkorn JK. Stroke: a family dilemma. Disabil Rehabil 1994; 16: 110-118.

156 Raffaele R, Rampello L, Vecchio I, Tornali C, Malaguarnera M. Trazodone therapy of the post-stroke depression. Arch Gerontol Geriatr 1996; 23 (Suppl 5): 217-220.

157 Robinson RG, Schultz SK, Castillo C, Kopel T, Kosier JT, Newman RM, Curdue K, Petracca G, Starkstein SE. Nortriptyline versus fluoxetine in the treatment of depression and in short-term recovery after stroke: a placebo-controlled, double-blind study. Am J Psychiatry 2000; 157: 351-359.

158 Bhakta BB, Cozens JA, Chamberlain MA, Bamford JM. Impact of botulinum toxin type A on disability and career burden due to arm spasticity after stroke: a randomised double blind placebo controlled trial. J Neurol Neurosurg Psychiatry 2000; 69: 217-221.

159 Bakheit AM, Pittock S, Moore AP, Wurker M, Otto S, Erbguth F, Coxon L. A randomized, double-blind, placebo-controlled study of the efficacy and safety of botulinum toxin type A in upper limb spasticity in patients with stroke. Eur J Neurol 2001; 8: 559-565.

160 Cifu DX, Stewart DG. Factors affecting functional outcome after stroke: a critical review of rehabilitation interventions. Arch Phys Med Rehabil 1999; 80: 35-39.

161 Hoenig H, Horner RD, Duncan PW, Clipp E, Hamilton B. New horizons in stroke rehabilitation research. J Rehabil Res Dev 1999; 36: 19-31.

162 Stroke Unit Trialists' Collaboration. Organised inpatient (stroke unit) care for stroke. In: The Cochrane Library, Issue 4. Oxford: Update Software, 2002.

163 Becker H, Stuifbergen AK, Sands D. Development of a scale to measure barriers to health promotion activities among persons with disabilities. Am J Health Promot 1991; 5: 449-454.

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

165 Whiteneck GG, Harrison-Felix CL, Mellick DC, Brooks CA, Charlifue SB, Gerhart KA. Quantifying environmental factors: a measure of physical, attitudinal, service, productivity, and policy barriers. Arch Phys Med Rehabil 2004; 85: 1324-1335.

166 Letts L, Law M, Rigby P, Cooper B, Stewart D, Strong S. Person-environment assessments in occupational therapy. Am J Occup Ther 1994; 48: 608-618.

167 Fougeyrollas P, Noreau L, Bergeron H, Cloutier R, Dion SA, St-Michel G. Social consequences of long term impairments and disabilities: conceptual approach and assessment of handicap. Int J Rehabil Res 1998; 21: 127-141.

168 Boldt C, Brach M, Grill E, Berthou A, Meister K, Scheuringer M, Stucki G. The ICF categories identified in nursing interventions administered to neurological patients with post-acute rehabilitation needs. Disabil Rehabil 2005; 27: 431-436.

169 Weigl M, Cieza A, Harder M, Geyh S, Amann E, Kostanjsek N, Stucki G. Linking osteoarthritis-specific health status measures to the International Classification of Functioning, Disability, and Health (ICF). Osteoarthritis Cartilage 2003; 11: 519-523.

170 Borchers M, Cieza A, Sigl T, Kollerits B, Kostanjsek N, Stucki G. Content comparison of osteoporosis-targeted health status measures in relation to the International Classification of Functioning, Disability and Health (ICF). Clin Rheumatol 2004; 14: 139-144

171 Stamm TA, Cieza A, Machold KP, Smolen JS, Stucki G. Content comparison of occupation-based instruments in adult rheumatology and musculoskeletal

rehabilitation based on the International Classification of Functioning, Disability and Health. Arthritis Rheum 2004; 51: 917-924.

172 Sigl T, Cieza A, Brockow T, Chatterji S, Kostanjsek N, Stucki G. Content comparison of low back pain-specific measures based on the International Classification of Functioning, Disability and Health (ICF). Clin J Pain 2006; 22: 147-153.

173 Cieza A, Stucki G. Content comparison of Health Related Quality of Life (HRQOL) instruments based on the International Classification of Functioning, Disability and Health (ICF). Quality of Life Research 2005; 14: 1225-1237.

174 American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for educational and psychological testing. Washington, DC: American Educational Research Association; 2004.

175 Cottrell RR, McKenzie JF. Health Promotion and Education Research Methods. Sudbury, MA: Jones and Bartlett Publishers; 2005.

176 Messick S. Meaning and Values in Test Validation: The Science and Ethics of Assessment. Educational Researcher 1989; 18: 5-11.

177 The Intercollegiate Working Party for Stroke. National clinical guidelines for stroke. Update 2002. London: Royal College of Physicians; 2002.

178 Scottish Intercollegiate Guidelines Network (SIGN). Management of patients with stroke. Rehabilitation, prevention and management of complications, and discharge planning. A national clinical guideline. Edinburgh: SIGN; 2002. (SIGN publication no. 64).

179 New Zealand Guidelines Group, Stroke Foundation of New Zealand Inc. Life after stroke. New Zealand guideline for management of stroke. Wellington: Stroke Foundation of New Zealand Inc; 2003.

180 European Stroke Initative Executive Committee, EUSI Writing Commitee. European Stroke Initiative Recommendations for Stroke Management – Update 2003. Cerebrovasc Dis 2003; 16: 311-337.

181 Diener E, Suh EM, Lucas RE, Smith HL. Subjective well-being: three decades of progress. Psychol Bull 1999; 125: 276-302.

182 Ferrans CE, Powers MJ. Psychometric assessment of the quality of life index. Res Nurs Health 1992; 15: 29-38.

183 Gilson BS, Gilson JS, Bergner M, Bobbit RA, Kressel S, Pollard WE, Vesselago M. The sickness impact profile. Development of an outcome measure of health care. Am J Public Health 1975; 65: 1304-1310.

184 Lai SM, Perera S, Duncan PW, Bode R. Physical and social functioning after stroke: comparison of the Stroke Impact Scale and Short Form-36. Stroke 2003; 34: 488-493.

185 Duncan PW, Lai SM, van Culin V, Huang L, Clausen D, Wallace D. Development of a comprehensive assessment toolbox for stroke. Clin Geriatr Med 1999; 15: 885-915.

186 Duncan PW, Horner RD, Reker DM, Samsa GP, Hoenig H, Hamilton B, LaClair BJ, Dudley TK. Adherence to postacute rehabilitation guidelines is associated with functional recovery in stroke. Stroke 2002; 33: 167-77.

187 Sacco RL, DeRosa JT, Haley EC Jr, Levin B, Ordronneau P, Phillips SJ, Rundek T, Snipes RG, Thompson JL; Glycine Antagonist in Neuroprotection Americas Investigators. Glycine antagonist in neuroprotection for patients with acute

stroke: GAIN Americas: a randomized controlled trial. JAMA 2001; 285: 1719-1728.

188 Petersen C, Morfeld M, Bullinger M. Testing and validation of the German version of the Stroke Impact Scale. Fortschr Neurol Psychiatr 2001; 69: 284-290.

189 Rankin J. Cerebral vascular accidents in patients over the age of 60. 2. Prognosis. Scott Med J 1957; 2: 200-215.

190 Linacre, JM. WINSTEPS Rasch measurement computer program. Chicago: Winsteps.com; 2005.

191 Wright BD, Masters GN. Rating Scale Analysis. Chicago: MESA; 1982.

192 Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989; 70: 857-860.

193 Smith RM, Schumacker RE, Bush JM. Using Item Mean Squares to Evaluate Fit to the Rasch Model. Journal of Outcome Measurement, 1998; 2: 66-78.

194 Linacre JM. Optimizing rating scale category effectiveness. J Appl Meas 2002; 3: 85-106.

195 Fisher WP. Reliability statistics. Rasch Measurement Transactions 1992; 6: 238.

196 Hagquist C, Andrich D. Is the Sense of Coherence-instrument applicable on adolescents? A latent trait analysis using Rasch-modelling. Pers Individ Dif 2004; 36: 955-68.

197 Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ 1995; 310: 170.

198 Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. Med Care 2000; 38 (9 Suppl): II28-42.

199 Bode RK, Lai JS, Cella D, Heinemann AW. Issues in the development of an item bank. Arch Phys Med Rehabil 2003; 84 (4 Suppl 2): S52-60.

200 Wright BD, Bell SR. Item Banks: What, Why, How. JEM 1984; 21: 331-45.

201 Kolominsky-Rabas PL, Sarti C, Heuschmann PU, Graf C, Siemonsen S, Neundoerfer B, Katalinic A, Lang E, Gassmann KG, von Stockert TR. A prospective community-based study of stroke in Germany--the Erlangen Stroke Project (ESPro): incidence and case fatality at 1, 3, and 12 months. Stroke 1998; 29: 2501-2506.

202 Stucki G, Ustun TB, Melvin J. Applying the ICF for the acute hospital and early post-acute rehabilitation facilities. Disabil Rehabil 2005; 27: 349-352.

203 Grill E, Ewert T, Chatterji S, Kostanjsek N, Stucki G. ICF Core Sets development for the acute hospital and early post-acute rehabilitation facilities. Disabil Rehabil 2005; 27: 361-366.

204 Dijkers MP, Kropp GC, Esper RM, Yavuzer G, Cullen N, Bakdalieh Y. Reporting on reliability and validity of outcome measures in medical rehabilitation research. Disabil Rehabil 2002; 24: 819-827.

205 Wright BD, Stone MH. Best Test Design. Chicago: Mesa; 1979.

206 Choppin B. Testing the questions - the Rasch model and item banking. In: Raggett MSJ, Tutt C, Raggett P, editors. Assessment and Testing of Reading: Problems and Practices. London: Ward Lock Educational; 1979.

X. APPENDIX

Appendix 1: The Comprehensive ICF Core Set for Stroke

ICF Code	ICF Category Title
Body Functions	
b110	Consciousness functions
b114	Orientation functions
b117	Intellectual functions
b126	Temperament and personality functions
b130	Energy and drive functions
b134	Sleep functions
b140	Attention functions
b144	Memory functions
b152	Emotional functions
b156	Perceptual functions
b164	Higher-level cognitive functions
b167	Mental functions of language
b172	Calculation functions
b176	Mental function of sequencing complex movements
b180	Experience of self and time functions
b210	Seeing functions
b215	Functions of structures adjoining the eye
b260	Proprioceptive function
b265	Touch function
b270	Sensory functions related to temperature and other stimuli
b280	Sensation of pain
b310	Voice functions
b320	Articulation functions
b330	Fluency and rhythm of speech functions
b410	Heart functions
b415	Blood vessel functions
b420	Blood pressure functions
b455	Exercise tolerance functions
b510	Ingestion functions
b525	Defecation functions
b620	Urination functions
b640	Sexual functions
b710	Mobility of joint functions
b715	Stability of joint functions
b730	Muscle power functions
b735	Muscle tone functions

ICF Code	ICF Category Title
b740	Muscle endurance functions
b750	Motor reflex functions
b755	Involuntary movement reaction functions
b760	Control of voluntary movement functions
b770	Gait pattern functions
Body Structures	
s110	Structure of brain
s410	Structure of cardiovascular system
s720	Structure of shoulder region
s730	Structure of upper extremity
s750	Structure of lower extremity
Activity and Participation	
d115	Listening
d155	Acquiring skills
d160	Focusing attention
d166	Reading
d170	Writing
d172	Calculating
d175	Solving problems
d210	Undertaking a single task
d220	Undertaking multiple tasks
d230	Carrying out daily routine
d240	Handling stress and other psychological demands
d310	Communicating with - receiving - spoken messages
d315	Communicating with - receiving - nonverbal messages
d325	Communicating with - receiving - written messages
d330	Speaking
d335	Producing nonverbal messages
d345	Writing messages
d350	Conversation
d360	Using communication devices and techniques
d410	Changing basic body position
d415	Maintaining a body position
d420	Transferring oneself
d430	Lifting and carrying objects
d440	Fine hand use
d445	Hand and arm use
d450	Walking
d455	Moving around
d460	Moving around in different locations
d465	Moving around using equipment
d470	Using transportation
d475	Driving

ICF Code	ICF Category Title
d510	Washing oneself
d520	Caring for body parts
d530	Toileting
d540	Dressing
d550	Eating
d570	Looking after one's health
d620	Acquisition of goods and services
d630	Preparing meals
d640	Doing housework
d710	Basic interpersonal interactions
d750	Informal social relationships
d760	Family relationships
d770	Intimate relationships
d845	Acquiring, keeping and terminating a job
d850	Remunerative employment
d855	Non-remunerative employment
d860	Basic economic transactions
d870	Economic self-sufficiency
d910	Community life
d920	Recreation and leisure
Environmental Factors	
e110	Products or substances for personal consumption
e115	Products and technology for personal use in daily living
e120	Products and technology for personal indoor and outdoor mobility and transportation
e125	Products and technology for communication
e135	Products and technology for employment
e150	Design, construction and building products and technology of buildings for public use
e155	Design, construction and building products and technology of buildings for private use
e165	Assets
e210	Physical geography
e310	Immediate family
e315	Extended family
e320	Friends
e325	Acquaintances, peers, colleagues, neighbours and community members
e340	Personal care providers and personal assistants
e355	Health professionals
e360	Health-related professionals
e410	Individual attitudes of immediate family members
e420	Individual attitudes of friends
e425	Individual attitudes of acquaintances, peers, colleagues, neighbours and community members
e440	Individual attitudes of personal care providers and personal assistants
e450	Individual attitudes of health professionals
e455	Individual attitudes of health-related professionals

ICF Code	ICF Category Title
e460	Societal attitudes
e515	Architecture and construction services, systems and policies
e525	Housing services, systems and policies
e535	Communication services, systems and policies
e540	Transportation services, systems and policies
e550	Legal services, systems and policies
e555	Associations and organizational services, systems and policies
e570	Social security services, systems and policies
e575	General social support services, systems and policies
e580	Health services, systems and policies
e590	Labour and employment services, systems and policies

Appendix 2: The Short ICF Core Set for Stroke

ICF Code	ICF Category Title
Body Functions	
b110	Consciousness functions
b114	Orientation functions
b730	Muscle power functions
b167	Mental functions of language
b140	Attention functions
b144	Memory functions
Body Structures	
s110	Structure of brain
s730	Structure of upper extremity
Activity and Participation	
d450	Walking
d330	Speaking
d530	Toileting
d550	Eating
d510	Washing oneself
d540	Dressing
d310	Communicating with - receiving - spoken messages
Environmental factors	
e310	Immediate family
e355	Health professionals
e580	Health services, systems and policies

Appendix 3: The Stroke Impact Scale 2.0 – Deutsche Version

Fragebogen über die Folgen eines Schlaganfalls (SIS) Deutsche Version

Zweck dieses Fragebogens ist eine Einschätzung der Folgen Ihres Schlaganfalls auf Ihre Gesundheit und Ihr Leben. Wir möchten wissen, wie sich Ihr Schlaganfall AUS IHRER PERSÖNLICHEN SICHT auf Sie ausgewirkt hat. Wir möchten Ihnen einige Fragen über mögliche Beeinträchtigungen stellen, die durch Ihren Schlaganfall verursacht wurden, und auch darüber, wie sich der Schlaganfall auf Ihre Lebensqualität ausgewirkt hat. Zum Schluss werden wir Sie bitten einzuschätzen, inwieweit Sie sich von Ihrem Schlaganfall erholt haben.

In den folgenden Fragen geht es um körperliche Probleme, die möglicherweise in Folge Ihres Schlaganfalls aufgetreten sind.

1. In der vergangenen Woche, wie viel Kraft hatten Sie Ihrer Meinung nach	**Sehr viel Kraft**	**Ziemlich viel Kraft**	**Etwas Kraft**	**Kaum Kraft**	**Gar keine Kraft**
a. im Arm, der am stärksten vom Schlaganfall betroffen war?	5	4	3	2	1
b. beim Zugreifen mit der Hand, die am stärksten vom Schlaganfall betroffen war?	5	4	3	2	1
c. im Bein, das am stärksten vom Schlaganfall betroffen war?	5	4	3	2	1
d. im Fuß/Knöchel, der am stärksten vom Schlaganfall betroffen war?	5	4	3	2	1

In den folgenden Fragen geht es um Ihr Gedächtnis und Ihr Denkvermögen.

2. In der vergangenen Woche, wie schwer ist es Ihnen gefallen....	**Gar nicht schwer**	**Etwas schwer**	**Ziemlich schwer**	**Sehr schwer**	**Außerordent-lich schwer**
a. sich an etwas zu erinnern, was man Ihnen gerade gesagt hat?	5	4	3	2	1
b. sich an Dinge zu erinnern, die am Vortag passiert sind?	5	4	3	2	1
c. sich daran zu erinnern, bestimmte Dinge zu tun (z.B. vereinbarte Termine wahrzu-nehmen oder Medikamente einzunehmen)?	5	4	3	2	1
d. sich an den aktuellen Wochentag zu erinnern?	5	4	3	2	1
e. Zahlen zusammenzuzählen und abzuziehen	5	4	3	2	1
f. sich zu konzentrieren?	5	4	3	2	1
g. schnell zu überlegen?	5	4	3	2	1
h. alltägliche Probleme zu lösen?	5	4	3	2	1

In den folgenden Fragen geht es um Ihr Befinden seit Ihrem Schlaganfall, um Veränderungen in Ihrer Stimmung und um die Fähigkeit, Ihre Gefühle zu beherrschen.

3. In der vergangenen Woche, wie oft...	Nie	Selten	Manchmal	Meistens	Immer
a. waren Sie traurig?	5	4	3	2	1
b. hatten Sie das Gefühl, dass es niemanden gibt, der Ihnen nahe steht?	5	4	3	2	1
c. hatten Sie das Gefühl, anderen eine Last zu sein?	5	4	3	2	1
d. hatten Sie das Gefühl, dass es nichts gibt, worauf Sie sich freuen können?	5	4	3	2	1
e. haben Sie sich wegen Fehlern, die Sie machten oder wegen Missgeschicken Vorwürfe gemacht?	5	4	3	2	1
f. haben Sie sich genauso über Dinge gefreut wie schon immer?	5	4	3	2	1
g. fühlen Sie sich nervös?	5	4	3	2	1
h. hatten Sie das Gefühl, das Leben sei lebenswert?	5	4	3	2	1
i. haben Sie mindestens einmal am Tag geschmunzelt und gelacht	5	4	3	2	1

In den folgenden Fragen geht es um Ihre Fähigkeit, sich anderen Menschen mitzuteilen und um Ihre Fähigkeit, Gelesenes oder bei einem Gespräch Gehörtes zu verstehen.

4. In der vergangenen Woche, wie schwer ist es Ihnen gefallen...	Gar nicht schwer	Etwas schwer	Ziemlich schwer	Sehr schwer	Außerordentlich schwer
a. den Namen eines Menschen zu nennen, der vor Ihnen stand?	5	4	3	2	1
b. zu verstehen, was Ihnen während einer Unterhaltung gesagt wurde?	5	4	3	2	1
c. auf Fragen zu antworten?	5	4	3	2	1
d. Gegenstände richtig zu benennen?	5	4	3	2	1
e. sich an einem Gespräch mit mehreren Leuten zu beteiligen?	5	4	3	2	1
f. ein Telefongespräch zu führen?	5	4	3	2	1
g. einen anderen Menschen anzurufen, einschließlich die richtige Telefonnummer zu finden und diese zu wählen?	5	4	3	2	1

In den folgenden Fragen geht es um Tätigkeiten, die möglicherweise zu Ihrem normalen Tagesablauf gehören.

5. In den vergangenen 2 Woche, wie schwer ist es Ihnen gefallen...	**Gar nicht schwer**	**Etwas schwer**	**Ziemlich schwer**	**Sehr schwer**	**Gar nicht möglich**
a. das Essen mit Messer und Gabel zu schneiden?	**5**	**4**	**3**	**2**	**1**
b. sich oben herum anzuziehen (von der Taille aufwärts)?	**5**	**4**	**3**	**2**	**1**
c. sich zu waschen (Bad, Dusche...)?	**5**	**4**	**3**	**2**	**1**
d. sich die Fußnägel zu schneiden?	**5**	**4**	**3**	**2**	**1**
e. schnell auf die Toilette zu kommen?	**5**	**4**	**3**	**2**	**1**
f. Ihre Blase zu kontrollieren (ohne Missgeschicke)?	**5**	**4**	**3**	**2**	**1**
g. Ihren Darm zu kontrollieren (ohne Missgeschicke)?	**5**	**4**	**3**	**2**	**1**
h. leichte Hausarbeiten zu erledigen?	**5**	**4**	**3**	**2**	**1**
i. einkaufen zu gehen?	**5**	**4**	**3**	**2**	**1**
j. mit Geld umzugehen (z.B. Wechselgeld richtig zurückgeben)?	**5**	**4**	**3**	**2**	**1**
k. sich um Ihre Geldangelegenheiten zu kümmern (z.B. Zahlung von monatlichen Rechnungen, Verwaltung des Girokontos)?	**5**	**4**	**3**	**2**	**1**
l. schwere Hausarbeiten zu erledigen?	**5**	**4**	**3**	**2**	**1**

In den folgenden Fragen geht es um Ihre Bewegungsfähigkeit (Mobilität) zu Hause und außer Haus.

6. In den vergangenen 2 Wochen, wie schwer ist es Ihnen gefallen...	**Gar nicht schwer**	**Etwas schwer**	**Ziemlich schwer**	**Sehr schwer**	**Gar nicht möglich**
a. zu sitzen, ohne das Gleichgewicht zu verlieren?	**5**	**4**	**3**	**2**	**1**
b. zu stehen, ohne das Gleichgewicht zu verlieren?	**5**	**4**	**3**	**2**	**1**
c. zu gehen, ohne das Gleichgewicht zu verlieren?	**5**	**4**	**3**	**2**	**1**
d. aus dem Bett auf einen Stuhl zu gelangen?	**5**	**4**	**3**	**2**	**1**
e. vom Stuhl aufzustehen, ohne sich mit den Händen abzustützen?	**5**	**4**	**3**	**2**	**1**
f. ungefähr 100 Meter weit zu Fuß zu gehen?	**5**	**4**	**3**	**2**	**1**
g. schnell zu gehen?	**5**	**4**	**3**	**2**	**1**
h. einen Treppenabsatz zu steigen?	**5**	**4**	**3**	**2**	**1**
i. mehrere Treppenabsätze zu steigen?	**5**	**4**	**3**	**2**	**1**
j. in ein Auto ein- und auszusteigen?	**5**	**4**	**3**	**2**	**1**

In den folgenden Fragen geht es um Ihre Fähigkeit, zum Gebrauch der Hand, die AM STÄRKSTEN von Ihrem Schlaganfall BETROFFEN war.

7. In den vergangenen 2 Wochen, wie schwer ist es Ihnen gefallen, die Hand, die am stärksten von Ihrem Schlaganfall betroffen war, zu benutzen, um...	**Gar nicht schwer**	**Etwas schwer**	**Ziemlich schwer**	**Sehr schwer**	**Gar nicht möglich**
a. schwere Sachen zu tragen?	**5**	**4**	**3**	**2**	**1**
b. einen Türknauf zu drehen?	**5**	**4**	**3**	**2**	**1**
c. eine Dose oder ein Glas zu öffnen?	**5**	**4**	**3**	**2**	**1**
d. Schnürsenkel zu binden?	**5**	**4**	**3**	**2**	**1**
e. eine kleine Münze aufzuheben?	**5**	**4**	**3**	**2**	**1**

In den folgenden Fragen geht es darum, wie sich Ihr Schlaganfall auf Ihre Fähigkeit ausgewirkt hat, die Dinge zu tun, an die Sie gewöhnt waren, die Ihnen wichtig sind und Ihrem Leben einen Sinn geben.

8. In den vergangenen 4 Wochen, wie oft waren Sie eingeschränkt in...	**Nie**	**Selten**	**Manchmal**	**Meistens**	**Immer**
a. beruflicher, ehrenamtlicher oder sonstiger Arbeit?	**5**	**4**	**3**	**2**	**1**
b. Ihren Aktivitäten mit anderen Menschen?	**5**	**4**	**3**	**2**	**1**
c. ruhigen Feizeitbeschäftigungen?	**5**	**4**	**3**	**2**	**1**
d. aktiven Freizeitbeschäftigungen?	**5**	**4**	**3**	**2**	**1**
e. Ihre Rolle als Familienmitglied oder als Freund/Freundin?	**5**	**4**	**3**	**2**	**1**
f. der Teilnahme an kirchlichen oder anderen religiösen Aktivitäten?	**5**	**4**	**3**	**2**	**1**
g. der Fähigkeit, Gefühle nahestehenden Personen gegenüber zu zeigen?	**5**	**4**	**3**	**2**	**1**
h. der Fähigkeit, Ihr Leben nach Ihren eigenen Wünschen zu bestimmen?	**5**	**4**	**3**	**2**	**1**
i. der Fähigkeit, anderen Menschen zu helfen?	**5**	**4**	**3**	**2**	**1**

CURRICULUM VITAE

PERSONAL DATA

Name:	Szilvia Geyh, née Békési
Date of Birth:	May 9th, 1974
Place of Birth:	Budapest, Hungary
Nationality:	Hungarian
Marital Status:	married

EDUCATION

1980 - 1986	Comprehensive school VÁCI UTCAI ÉNEK-ZENEI ÁLTALÁNOS ISKOLA, BUDAPEST
1987 - 1994	Abitur GYMNASIUM OF CISTERCIANS SELIGENTHAL, LANDSHUT
1994 - 2001	Diploma in Psychology CATHOLIC UNIVERSITY EICHSTAETT
2003 - 2005	Master of Public Health (postgrad.) LUDWIG-MAXIMILIAN UNIVERSITY, MUNICH

PROFESSIONAL EXPERIENCE

1997	Internship Department of Psychosomatics CHILDREN'S HOSPITAL ST. MARIEN, LANDSHUT
1998	Internship Department of Neuropsychology M & I CLINIC FOR PHYSICAL MEDICINE AND MEDICAL REHABILITATION, ICHENHAUSEN
2001 to date	Research scientist ICF RESEARCH BRANCH, WHO COLLABORATING CENTER FOR THE FAMILY OF INTERNATIONAL CLASSIFICATIONS, GERMANY (DIMDI), INSTITUTE FOR HEALTH AND REHABILITATION SCIENCES, LUDWIG-MAXIMILIAN UNIVERSITY, MUNICH

FURTHER EDUCATION

2004	Instrument Design with Rasch IRT. Correspondence Course. MURDOCH UNIVERSITY, PERTH, WESTERN AUSTRALIA.
2006	Practical Rasch Measurement. Online Course. STATISTICS.COM, ARLINGTON, USA.

PUBLICATIONS

Geyh S, Wendel C, Heel S, Fries W. *Kognitive Funktionen und Selbsteinschätzung der Alltagskompetenz in der ambulanten neurologisch-neuropsychologischen Rehabilitation.* Zeitschrift für Neuropsychologie 2002;13: 281-289.

Weigl M, Cieza A, Harder M, Geyh S, Amann E, Kostanjsek N, Stucki G. *Linking osteoarthritis-specific health-status measures to the International Classification of Functioning, Disability and Health (ICF).* Osteoarthr Cartil 2003;11:519-523.

Brockow T, Duddeck K, Geyh S, Schwarzkopf S, Weigl M, Franke T, Brach M. *Identifying the concepts contained in outcome measures of clinical trials on breast cancer using the International Classification of Functioning, Disability and Health as a reference.* J Rehabil Med 2004;Suppl 44:43-48.

Brockow T, Wohlfahrt K, Hillert A, Geyh S, Weigl M, Franke T, Resch KL, Cieza A. *Identifying the concepts contained in outcome measures of clinical trials on depressive disorders using the International Classification of Functioning, Disability and Health as a reference.* J Rehabil Med 2004;Suppl 44:49-55.

Geyh S, Kurt T, Brockow T, Cieza A, Ewert T, Omar Z, Resch KL. *Identifying the concepts contained in outcome measures of clinical trials on stroke using the International Classification of Functioning, Disability and Health as a reference.* J Rehabil Med 2004;44 Suppl:56-62.

Stucki G, Cieza A, Geyh S, Battistella L, Lloyd J, Symmons D, Kostanjsek N, Schouten J. *ICF Core Sets for rheumatoid arthritis.* J Rehabil Med 2004;Suppl 44:87-93.

Cieza A, Stucki A, Geyh S, Berteanu M, Quittan M, Simon A, Kostanjsek N, Stucki G, Walsh N. *ICF Core Sets for chronic ischaemic heart disease.* J Rehabil Med 2004;Suppl 44:94-9.

Geyh S, Cieza A, Schouten J, Dickson H, Frommelt P, Omar Z, Kostanjsek N, Ring H, Stucki G. ICF Core Sets for stroke. J Rehabil Med 2004;44 Suppl:135-41.

Cieza A, Geyh S, Chatterji S, Kostanjsek N, Üstün B, Stucki G. *ICF linking rules: an update based on lessons learned.* J Rehabil Med 2005;37:212-218.

Ewert T, Geyh S, Grill E, Cieza A, Zaisserer S, Stucki G. *Die Anwendung der ICF in der Neurorehabilitation anhand des ICF-Modellblattes und der ICF Core-Sets.* Neurol Rehabil 2005;11:179–188.

Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun BT, Stucki G. *Identification of candidate categories of the International Classification of Functioning Disability and Health (ICF) for a Generic ICF Core Set based on regression modelling.* BMC Med Res Methodol. 2006 Jul 27;6(1):36 [Epub ahead of print].

Stamm T, Geyh S, Cieza A, Machold K, Kollerits B, Kloppenburg M, Smolen J, Stucki G. *Measuring functioning in patients with hand osteoarthritis--content comparison of questionnaires based on the International Classification of Functioning, Disability and Health (ICF)*. Rheumatology (Oxford). 2006 May 2; [Epub ahead of print].