FACULTY OF HEALTH AND MEDICAL SCIENCES
UNIVERSITY OF COPENHAGEN

PhD thesis

Tina Hansen, OT, MSc.OT

Dysphagia in frail elderly patients, from an occupational therapy perspective

Danish translation and validation of the McGill Ingestive Skills Assessment for observation-based measurement of occupational performance in eating and drinking activities

Dysphagia in frail elderly patients, from an occupational therapy perspective

Danish translation and validation of the McGill Ingestive Skills Assessment for observation-based measurement of occupational performance in eating and drinking activities.

Tina Hansen, OT, MSc.OT

PhD thesis

Dysphagia in frail elderly patients, from an occupational therapy perspective: Danish translation and validation of the McGill Ingestive Skills Assessment for observation-based measurement of occupational performance in eating and drinking activities.

Dysfagi hos skrøbelige ældre patienter, set fra et ergoterapeutisk perspektiv: Dansk oversættelse og validering af McGill Ingestive Skills Assessment til observationsbaserede måling af aktivitetsudførelse i spise og drikke aktiviteter.

PhD thesis submitted June 2012

Public defense at Herlev University Hospital May 6, 2013

Author: Tina Hansen, OT, MSc.OT, Department of Occupational Therapy, Herlev University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark.

Official opponents

- Professor, MD, DMSci. Peter Schwarz (Chairman), Faculty of Health Science, Copenhagen University, Copenhagen, Denmark; and Research Center of Aging and Osteoporosis, Department of Medicine, Glostrup University Hospital, Denmark.

- Associate professor, MD, PhD, John Brodersen Research Unit and Section for General Practice, Department of Public Health, University of Copenhagen, Copenhagen, Denmark.

- Professor, PT, PhD, Liv Inger Strand, Physiotherapy Research Group, Department of Public Health and Primary Health Care, University of Bergen, Norway.

Scientific Advisors

- Faculty supervisor Jens Faber, Professor, MD, DMSci, Department of Medicine/Endocrinology, Herlev University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark. Department of Orthopaedics and Internal Medicine at the University of Copenhagen, Denmark.

- Hanne Trine Sander Pedersen, MD. Department of Medicine/Geriatrics, Herlev University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark.

- Heather Christine Lambert, BSc.OT, PhD, School of Physical and Occupational Therapy McGill University Davis House, 3654 Promenade Sir William Osler Montreal, Quebec H3G 1YS, Canada.

Contents

Abbreviations

AD index	Average deviation index
ANOVA	Analysis of variance
BI	Barthel-100 index
CI	Confidence interval
CTT	Classical test theory
CVI	Content validity index
CVQ	Content validity questionnaire
df	Degrees of freedom
DIF	Differential item functioning
ICC	Intraclass correlation coefficient
ICF	International Classification of Functioning, Disability and Health
Kw	Weighted Kappa
LID	Local item dependency
LOA	Limits of agreement
MISA	McGill Ingestive Skills Assessment, Canadian version
MISA-DK	McGill Ingestive Skills Assessment, Danish version
MMSE	Mini-Mental Status Examination
NOT-S	Nordic Orofacial Test-Screening
PO	Percentage of observed agreement
PROM	Patient reported outcome measure
PSI	Person-Separation index
SD	Standard deviation
SDC	Smallest detectable change
SEM	Standard error of measurement
VFS	Videofluoroscopy
WHO-5	WHO-five, Well-Being index
WST	Water swallow test

List of publications

1. Hansen T, Lambert HC, Faber J. Content validation of a Danish version of "The McGill Ingestive Skills Assessment" for dysphagia management. Scand J Occup Ther 2011;18(4):282-293.

2. Hansen T, Lambert HC, Faber J. Validation of the Danish version of the McGill Ingestive Skills Assessment using classical test theory and the Rasch model. Disabil Rehabil 2012;34(10):859-868.

3. Hansen T. Lambert HC, Faber J. Reliability of the Danish version of the McGill Ingestive Skills Assessment for observation-based measures during meals. Scand J Occup Ther 2012;19(6):488-496.

1. Introduction

Elders form an increasing proportion of the hospital population due to demographic aging in Denmark (1,2). The frail elderly patient is particularly vulnerable because of decreased physiological reserves, high prevalence of chronic diseases and comorbidity (3,4). This often causes dysphagia; i.e., difficulty in swallowing (5-8). Dysphagia may result in aspiration pneumonia (5,9-11), malnutrition and dehydration (5,10,11), and is associated with increased morbidity and mortality (5,10, 11), increased length of hospital stay (10-12), and discharging to institutional care (12,13). In addition, as eating and drinking is involved in many facets of a person's daily life and is a form of social interaction (14,15), dysphagia may also lead to social isolation and decreased quality of life (15-21). The dysphagic patient's ability to eat and drink safely, efficiently, independently and with pleasure during meals is an important focus in dysphagia management (5,14,22,23). Measurement instruments[1] with this focus and evidence of psychometric properties in terms of validity and reliability are needed for clinical practise and research (22). The Canadian" McGill Ingestive Skills Assessment" (MISA) (24) fulfils these requirements (25). The purpose of the MISA is to measure frail elderly patients' ability to ingest a variety of food and liquids safely, efficiently and independently during a meal, and is intended to be used in treatment planning and outcome measuring (24). In this thesis, the translation and validation of a Danish version of the MISA (MISA-DK) is addressed.

2. Background

2.1. Dysphagia in the older population

Age-related changes in the swallowing mechanism of otherwise healthy elders (presbyphagia) is manifested by sarcopenia and changes in sensorimotor acuity and efficiency, which decreases the tempo, flexibility and strength of the structures for eating, drinking and swallowing (5,7,22,26). This reduced functional reserve increases the risk of dysphagia under stressful conditions (5-8,22, 26). Prevalence estimates of dysphagia among independent living elders older than 65 years are in the range of 11% to 33% (27-29). The most common diseases of aging associated with dysphagia are neurological, neuromuscular and structural disorders (5-8,10,22). Furthermore, seemingly unre-

[1]In paper I-III, the terms "assessment", "clinical assessment", "assessment instruments", "measurement instrument", "clinical measurements" and "assessment tool" are used interchangeably. Throughout this summary, measurement instrument will be used. Measurements are the data obtained by measuring in order to ascertain the dimension, quantity or capacity of a latent trait variable. This includes the application of a standard scale, thus translating direct observations/patient reports to a numerical scoring system. Assessment is the overall process of collecting information and includes multiple data-collecting instruments and sources of information (42).

lated conditions such as frailty (11,13,30) have been found to cause dysphagia in hospitalised elders. Frailty is characterised by a multisystem reduction in physiological capacity leaving the individual with increased risk of diseases and disability (3,4). Malnutrition and sarcopenia are core features of frailty (3,4), and the underlying factors involved in the development of dysphagia in old age and frailty is assumed to be interrelated (5,30). The prevalence of dysphagia among frail elders is estimated to approximately 29% when acutely admitted to geriatric care (13), 55% when hospitalized with pneumonia (12), 54% when residing in the community and receiving healthcare (31) and 51% when residing in a nursing home (32).

2.2. The dysphagia assessment process

To capture the true impact of dysphagia and its interventions, it is suggested that the assessment process includes information based on all the components of the International Classification of Functioning, Disability and Health (ICF) (23). ICF provides a framework, in which functioning is described as the complex interplay of the health components *body functions* (the physiological functions of body systems), *body structures* (the anatomical parts of the body), *activity* (the execution of a task or action) and *participation* (involvement in a life situation), and the contextual factors: *environmental factors* (physical, social and attitudinal milieu) and *personal factors* (particular circumstances in a person's life and living) (33). The ICF as a classification represents a catalogue of mutually exclusive ICF categories that refer to each component (33).The above cited prevalence studies used different measurement instruments such as self-reported dysphagic symptoms, swallowing trials with few different liquid and food consistencies, clinical examination or videofluoroscopy (VFS) (12,13,19,27-32). All of these are well-recognised within dysphagia management (5-7,22, 34), but they addresses predominantly the pharyngeal aspects of swallowing, which are related to the ICF components for body structures and body functions (23). Given that these instruments are performed in an artificial environment, they do not reflect the complexity of eating and drinking in a natural context and may lead to recommendations with limited relevance for a given patient (22,23,35). Hence, information based on the ICF components for activity, participation and contextual factors ought to be implemented in clinical practise and research (22,23). These ICF components relates very closely to occupational performance of the patient, which is the domain of concern in occupational therapy (36-38). Occupational performance refers to the complex and dynamic interaction of the physical, cognitive and affective performance components within the individual, the occupation and the physical, social and cultural environment. Occupations are those purposeful, meaningful and cultural significant activities that people do in their daily lives in order to develop

and maintain health and well-being (36-38).

2.3. Dysphagia from an occupational therapy perspective

In Denmark, occupational therapists have a key role in the multidisciplinary dysphagia management (39), and they consider swallowing as an integral part of occupational performance in eating and drinking activities in a natural context (14,38,40,41). This is conceptualised as four interdependent phases (14,38,41), which require coordinated cognitive and sensorimotor activity involving both subcortical and cortical cerebral areas (22).

1. The pre-oral phase is aimed for anticipation to the meal. It comprises sensation, perception, and cognition related to visual, tactile, and olfactory inputs, physiological factors and the skills of self-feeding with the coordination of the movements of the eyes, arms and hands together with the movements of the trunk, head and jaw (14,22,35,38,41).
2. The oral phase is aimed for bolus formation and propulsion. It comprises sensation, saliva secretion, motor planning, and jaw, labial, buccal and tongue muscle tone, movement and coordination (14,22,41).
3. The pharyngeal phase is aimed at swallowing (i.e. bolus propulsion), which is triggered by an activation of pharyngeal mechanoreceptors that send information to the brainstem swallowing centre containing the central pattern generator. Synchronously, a centrally generated respiration pause occurs. The phase comprises closure of the nasopharynx (soft palate elevation), tongue base movement, closure of the airway (hyoid bone elevation and anterior displacement, lowered epiglottis, and vocal cords closure), and contraction of pharyngeal constrictors and opening of the upper oesophageal sphincter muscles. (14,22,41).
4. The oesophageal phase begins with the opening of the upper oesophageal sphincter, which is followed by oesophageal peristalsis transporting the bolus into the stomach (14,22,41).

Difficulties in any of these phases may impair the efficiency and safety of the process of swallowing (5,14,22,35,41) and influences on the patient's occupational performance in eating and drinking activities (14,40). In their assessment process, occupational therapists implements performance analysis, which involves observation of the quality of the patient's occupational performance (37, 42). This necessitates measurement instruments with evidence of validity and reliability (36,42). Overall, validity is the degree to which a measurement instrument measures the construct(s) it purports to measure, and reliability is the degree to which scores for patients who have not changed are the same for repeated measurements under several conditions (36,42-45).

Within Danish occupational therapy, one dysphagia specific measurement instrument is available

and contains four different assessment forms of which one addresses the patient's occupational performance in eating and drinking: "Screening of oral ingestion" (41). However, the assessment form does not provide a specific rating scale. Furthermore this instrument has no documented evidence on validity or reliability. Therefore, it is neither compatible with an evidence-based practise nor appropriate for use in research (36,42-46). In recognition of this, we undertook a literature review in order to identify valid and reliable instruments suitable for measuring elderly dysphagic patients' occupational performance in eating and drinking activities (25). A search in CINAHL, Pubmed, PsycINFO, and Web of Science using combinations of several search terms related to the area, related citations and references from retrieved papers resulted in identification of 14 measurement instruments. Of these, eight converged with the conceptualisation of occupational performance in eating and drinking activities within Danish occupational therapy (39,41), and formed a scale of items associated to the pre-oral, the oral and the pharyngeal phases. The oesophageal phase was not considered, since impairment in this area requires diagnostics by the medical profession (7,8,22). The evidence of validity and reliability of the measurement instruments was quality appraised using predefined criteria related to the sample size, the used statistics and the magnitudes of the validity and reliability estimates (36,43,44). Reliability was not documented for three, was poor for three and adequate for two. Validity was poor for four, adequate for three and excellent for one (25). To be used by occupational therapists, only "the McGill Ingestive Skills Assessment" (MISA) (24) exhibited adequate evidence of both validity and reliability (25,47-49).

2.4. The McGill Ingestive Skills Assessment (MISA)

The MISA is administered as observation during a natural meal, which is planned together with the patient taking individual food preferences or dietary restrictions into account. In the MISA, the conceptualisation of occupational performance in eating and drinking activities is based on a construct of skills termed "ingestion" (35). These are described in 43 ingestive skill items relating to observable actions necessary to complete a meal efficiently and safely. The items are distributed into six subscales: positioning (4 items) addressing the patient's ability to maintain a position that is safe for eating and drinking; self-feeding skills (7 items) addressing the patient's self-feeding skills, behaviour, and judgment; liquid ingestion (7 items) addressing the patient's oropharyngeal skills for liquids; solid ingestion (12 items) addressing the patient's oropharyngeal skills for solids; and texture management addressing the patient's capability to willingly and safely management of solid food (8 items) and liquid textures (5 items). Each item is scored on a 3-point ordinal scale (1= absent of ingestive skill, 2= insufficient ingestive skill performance, 3= adequate ingestive skill perfor-

mance), which is summarized to give subscale scores and a total score. The scores are documented at a four-page score sheet. The interpretation of the observation and scoring is to be supported by specific item- and score descriptions in the instruction manual (24).

2.5. Using measurement instruments in different cultures

The MISA was developed in English by Canadian occupational therapists. Therefore, functional equivalence has to be considered when it is to be used in a Danish occupational therapy context (43,50). Like this, uniform administration and interpretation across different languages and cultures can be assured allowing comparison of outcome results across borders and in multinational research (43,50-52). Functional equivalence concerns the extent to which a measurement instrument does what it is supposed to do equally well in two or more cultures (43,50). Thus, it ought to be ensured that the MISA-DK is equivalent in terms of: a) conceptualisation of the construct, b) relevancy and appropriateness of the items, c) semantics, d) relevancy and appropriateness of measurement methods, and d) the psychometrical properties (43,50). Functional equivalence is to a great extent about construct validity (53). The view of construct validity prevailing today, subsumes all validity and reliability aspects under one unified validity concept, and is concerned with an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on measurement scores (54-56). The unified validity concept includes five sources of evidence (54,56), and is briefly outlined below in conjunction with definitions and results of the psychometrical properties addressed for the MISA.

Test content: address whether the content, items, subscales and formats of the measurement instrument is an adequate and representative reflection of the construct to be measured, and it includes the traditional notion of content validity (54-56). The MISA was developed via an extensive literature review, focus-group methodology and pilot testing (47).

Response processes: address whether the responses of the persons do fit the intended construct being measured (54-56). It includes the traditional examination of flour and ceiling effects (54) and whether the response pattern to the items fit the intended defined construct (54-56). For the MISA, the ordinal structure of the score categories for each item was evaluated using expert judgments as well as statistical analyses of the score distributions, which resulted in clarifications of some score descriptors before its publishing (24,47).

Internal structure: address whether the relationships between the items match the intended construct being measured and whether measurements are generalizable and reliable, and it includes evaluat-

ing dimensionality of scale items (54-56). For the MISA, the assignation of items into subscales was driven by theory and by examining inter-item and item-scale correlations (47). Internal structure also includes the traditional categories of reliability (54), such as: a) interrater: the degree to which scores for the same patient who have not changed are the same when measured by different raters on the same occasion; b) intrarater: the degree to which scores for the same patient who have not changed are the same when measured by the same rater on different occasions, and c) internal consistency: the degree of interrelatedness among the items (36,42-45). For the MISA, analyses using the Intraclass correlation coefficient (ICC) revealed good to excellent inter- and intra-rater reliabilities for the subscales and excellent for the total scale (48). The internal consistency was found adequate by means of Cronbach's alpha (α) (24).

Relations to other variables: address whether there are relationships between measurement scores and other variables to which they are expected to correlate with or predict (54-56). It includes the traditional categories of construct validity (54), such as: a) convergent validity; the degree to which the measurement scores are related to scores of other instruments measuring theoretical related constructs, and b) known-groups validity; the degree to which the measurement instrument demonstrates different scores for groups known to vary on the construct being measured; and criterion validity, such as c) predictive validity; the degree to which the measurement scores predicts specific future events (36,42-45). The MISA total score correlated significantly to measurement scores of physical function and cognition (convergent validity) and it discriminated significantly among patients wearing dentures versus those who did not (known-groups validity) (48). In addition, decreasing MISA total scores increased the risk of death (predictive validity) (49).

Consequences of testing: address whether anticipated or unanticipated negative or positive effects occur and includes analysis of false positive and negative results (i.e. sensitivity and specificity) (54-56). This has not been addressed for the MISA.

2.6. Methods for establishing functional equivalence

Since a crucial part of functional equivalence is the translation, it is recognized that a comprehensive approach involving many steps is needed (50,51,57,58). In this process, the content validity of the translation is evaluated in the target group as well (51), i.e. Danish occupational therapists in the case of MISA. If the content validity is not sufficient, it might be necessary to revise, add or remove items (59,60). Additionally, in relation to content validity, it is increasingly recognized that by linking a measurement instrument to the ICF (33), further information and understanding of the measurement are provided (61,62).

The statistical methods for testing the equivalence of the psychometrical properties can be based on two theories: Classical Test Theory (CTT) and Item Response Theory (IRT) (43,52-56,63,64). The MISA was developed and validated based on CTT expressing a linear association that the observed score contains a true score and an error score (43,56,63,64). In order to establish the equivalence of the psychometric properties of the MISA-DK, investigation of the convergent and known-groups validity, the internal consistency reliability using Cronbach's alpha and inter- and intrarater reliability using the ICC is relevant (50,51). However, it is worth noticing, that due to sample dependency of the statistical methods within CTT (43-45,63-66) it might be unrealistic to expect similar results (50). In addition, as the ICC measures relative reliability with an estimate between 0 and 1 (43-45, 65), it might be appropriate to investigate the absolute reliability (45,65-69) of MISA-DK as well. Opposite to the ICC, absolute reliability estimates are population independent, are expressed in the actual units of a measurement, and provides information on the absolute measurement error connected with an individual's measurement score (45,65-69).

Since the items in the MISA, and thus MISA-DK, are intended to be summarized into subscale- and total scores and the methods within CTT predominantly focus on test-level statistic, unidimensionality must be demonstrated (46,56,64,70-73). This has not been addressed for the MISA (47). Methods for examining the dimensionality of a measurement instrument are factor analysis derived from CTT (43,44,46,56,63) or IRT models (43,56,63,64). IRT models have some advantages over factor analytic methods, such as: sample independency; it is not necessary to assume normal distribution of the data; and information from the response patterns is analysed as opposed to the more limited information from correlation matrices used in factor analysis (72,74-76). IRT is a group of models for expressing the association between observed (actual) item performance and the underlying ability (unobserved) or latent trait (43,63,64,72). This association is described via S-shaped item characteristic curves, which are non-linear and monotonic (43,56,63,64,72). The IRT is based on two basic assumptions: 1) the latent trait variable is a continuous unidimensional construct that explains the covariance among the item responses; and 2) the item responses are conditionally independent of one another given the latent trait variable (43,63).
Depending on the specific IRT model, each item is characterized by one or more model parameters: 1) an item difficulty or threshold parameter, which is the point on the latent trait variable where a patient has 50% chance of succeeding an item; 2) an item discrimination parameter, which indicates how well an item discriminates between patients below and above the threshold parameter, and 3) a

pseudo guessing parameter, which accounts for the performance of low-ability examinees on multiple choice items (43,63,64). Of the different IRT models, only the Rasch model (77) comply with the requirements of fundamental measurement in terms of specific objectivity, which implies invariance (64,71-74,77-79). That is, comparison between any two patients should be independent of which items of the measurement instrument are used, and vice versa (55,64,72-74,77-79). Another important property of the Rasch model is that the total score is a sufficient statistic (55,64,77,78). That is, all available information is in the patient's or the item's total score and no information on the response pattern is needed (64,77,78).

The Rasch model was originally developed as a one-parameter logistic model for dichotomous response options and it includes only the item difficulty parameter (i.e. all of the items have equal discriminating ability) (64,71,77-79). The theoretical background of the Rasch model (and IRT models) is a Guttman scale (43,56,73,79). A hypothetically deterministic Guttman scale consists of a unidimensional set of items, which are ranked in order from least to most difficult. For any total score, the pattern of responses can be inferred (43,56,79). The Rasch model is a probabilistic counterpart of the Guttman scale and specifies that the probability of a patient succeeding an item is a logistic function of the difference between the patient's ability level and the difficulty of the item (43,56,64,73,77-79). The logistic formulation gives linearization of the probabilities (log-odd units) (56,70,74,77,79). This allows that the estimated item and person parameters are placed on the same logit-scale, which is expressed in equal-interval units and is centred by a mean item location of zero. As such, the patient has a true ability score (location) on a continuous latent variable of less or more (43,56,64,73,77,79), e.g., ingestive skill abilities. In the case of ordered categories, such as in the MISA-DK, there are generalizations of the Rasch model (56,64,79-81): the rating scale model, which assumes a common rating scale structure acorns all items (80); and the partial credit model, which assumes that each item has its own rating scale structure (81). In these models, threshold parameters are included, which refers to the point between two adjacent score categories where either score is equally probable, and monotonicity is expected (56,64,80-83).

Whether the generated item and person parameters are valid, i.e. show criterion-related construct validity (78,84), depends on how well the data fits the assumptions in the Rasch model in terms of unidimensionality, monotonicity, local item independence and invariance (43,56,63,64,71-73,77-79). Invariance implies that the hierarchical order of items should remain the same at different abil-

ity levels and that the items do not present differential item functioning (DIF); i.e. patients who have equal ability levels may not have different probabilities of succeeding an item because of e.g. age or gender (63,64,73,77-79). Using the Rasch model to review the psychometric properties of MISA-DK will provide further information on the items and their category structure as well as whether a summation of the items into a total score and subscale scores can be justified (64,70-73, 79). Such information might aid in the interpretation of the undertaken analysis within CTT, direct further analysis within CTT and provide the means to improve the validity of MISA-DK (64,73).

3. Hypothesis and aims for the thesis

The MISA operationalizes occupational performance in eating and drinking as observable ingestive skills and has documented acceptable psychometric properties. It is, therefore, obvious to formulate the hypothesis that the MISA can be used by occupational therapists in a Danish context, i.e. has functional equivalence and can serve as an instrument for measuring dysphagia from an occupational therapy perspective.

3.1. Overall aim

The overall aim was to produce a functional equivalent Danish version of the MISA, which possess adequate levels of validity and reliability.

3.2. Specific study-related aims

- To translate and cultural adapt the MISA into the Danish MISA-DK, and to evaluate the content validity the MISA-DK by expert-panel judgment, pilot-testing and content identification and quantification using the ICF as a frame of reference (Study I).
- To establish equivalence of the psychometrical properties of MISA-DK with regard to the internal consistency reliability, the convergent validity and known-groups validity, and to extend the evaluation of the validity using Rasch analysis (Study II and supplementary analyses of data in study II presented in this summary).
- To establish equivalence of the psychometrical properties of MISA-DK with regard to the relative inter- and intra-rater reliability, and to extend the evaluation of the reproducibility of the MISA-DK in terms of absolute reliability and item level reliability (Study III and supplementary analyses of data in study III presented in this summary).

4. Methods

4.1. Study design

The work was initiated with translation and cultural adaptation of the MISA into the MISA-DK (Study I) of which, the psychometric properties were evaluated (Study II-III). The studies were all empirical (Table I), and the data were collected primo-2009 to mid-2011 and were analysed quantitatively.

Table I. Overview of the thesis and the characteristics of the studies underpinning Paper I-III.

	Design	Psychometric properties	Participants	Data collection
Study I	Translation. Prospective questionnaire based survey	Content validity.	Danish occupational therapists (n=23)	Self-administrated Questionnaire
Study II	Prospective, consecutive, cross-sectional	Internal consistency reliability. Known-groups validity. Convergent validity. Criterion-related construct validity[2].	Danish patients (n=110)	In-person observation using MISA-DK and data on external validation variables
Study III	Cross-sectional two-rater and test-retest design	Relative inter- and intra-rater reliability. Absolute inter- and intra-rater reliability.	Danish occupational therapists (n= 38) Danish patients (n=102)	Video-based observation using MISA-DK

4.2. Study population

The study population in this thesis included Danish occupational therapists (study I and III), and acutely hospitalised Danish patients (study II-III).

Study I and III - Occupational therapists:

For study I, twenty-one occupational therapists were recruited from main hospitals and rehabilitation centres at the Zealand of Denmark; thirteen participated as content-experts and 16 participated as pilot-testers, of which 8 also participated as content-experts. All were experienced in dysphagia management (average range in years: content-experts, 2 -17 years; pilot-testers, 1-17 years) and about 50% had specialized post-education in dysphagia. Additionally, two occupational therapists were recruited from the Danish ICF network (85) and participated as ICF-experts in the content identification of the MISA-DK.

For study III, a total of 38 occupational therapists participated as raters and were recruited from main hospitals and rehabilitation centres at the Zealand of Denmark. All were experienced in dys-

[2] In paper II, Criterion-related construct validity is referred to as internal construct validity

phagia management (average range in years, 0.5 -17 years) and about 45% had specialized post-education in dysphagia.

Study II and III - Patients:

Patients were recruited from two general medical wards at Herlev hospital in the Capital Region of Copenhagen. The patients were consecutively included within 48 hours of admission if they were over 65 years, were not terminally ill, would require more than 2 days of hospitalization and were able to give personal information and written informed consent. The patients were excluded if they did not fulfil five criteria for direct swallowing measurement (41), namely the ability to: remain alert for at least 15 minutes, sit in a chair or bed in at least a 60° upright position, swallow saliva, cough voluntarily and clear the throat twice.

The inclusion was performed by the author (TH) from December 2009 to February 2011. Of 439 eligible patients, 168 were unable to give personal information and written informed consent and 87 declined. Of the remaining 184 patients, 74 (40%) were unable to fulfil on or more of the swallowing criteria. This resulted in the inclusion of 110 patients, of which all agreed upon participation.

The sample was represented by 50% males and females, respectively. The mean age was 81.9 (SD 7.6) years. The patients had on average 2.15 admission diagnoses (SD 1.1) and on average 2.7 chronic medical conditions (SD 1.6). The main diagnostic characteristics were distributed as follows: 63% had diseases of the circulatory system and 25% had sequelae after stroke, 57% had diseases of the respiratory system (chronic obstructive lung disease and/or asthma), 44% had a diagnosis of pneumonia, 35% had diseases of the musculoskeletal system, 25% had diabetes mellitus, 16% had urinary tract infection and 10% had diseases of the nervous system such as Parkinson’s disease or epilepsy.

All 110 patients were included for the analysis in study II. Of the 110 patients, 102 agreed to be video-recorded during a meal and were included for the analysis of the inter- and intra-rater reliability in study III.

4.3. Instrumentation and procedures

In Study I, permission to translate the MISA was attained from Heather C. Lambert (the primary author of MISA) and the Canadian Association of Occupational Therapists from whom TH has bought the Copyright on the MISA-DK (86). The translation procedure used a collaborative approach in four phases (51,57), and Heather C. Lambert was continuously involved.

The MISA was translated into the MISA-DK by three translators (two certified translators and one bilingual occupational therapist), who made independent parallel translations. Secondly, a bilingual review committee of two occupational therapists, a dietician, and TH examined the semantic equivalence of all translations and the MISA in terms of comprehensiveness, accuracy, cultural relevance, linguistic quality and naturalness (50); decided on the most appropriate translations and produced a synthesis of these. One of the certified translators was occasionally requested to retranslate problematic items and sections. Thirdly, a bilingual occupational therapist who is a native speaker of English (USA) performed a thorough evaluation of the semantic equivalence between the consensus version and the MISA. This resulted in a contemporary version of the MISA-DK drawn up by TH. In the fourth phase, the MISA-DK was content validated using content experts (59,60,87, 88). The content experts were introduced to the MISA-DK via a two hour introduction meeting and responded independently to a Content validity questionnaire (CVQ) within three weeks. The CVQ used a scoring system proposed by Lynn (88), and included six content validity domains: adequacy of the item terms in reflecting the item content, clarity of the item descriptors, clarity of the score descriptors, relevancy of each item, clarity of all the paragraphs in the instruction manual and all the sections of the score sheet. For each content validity domain, a four-point Likert scale was used: 1 = not at all adequate/clear/relevant, 2 = needs major modifications to be adequate/clear/ relevant, 3 = needs minor modifications to be adequate/clear/relevant, 4 = very adequate/clear /relevant. Based on these judgements, subsequent discussions with the experts at a follow-up meeting and dialogue with Heather C. Lambert, a final version of the MISA-DK was produced by TH. Due to the Copyright agreement on the MISA (86), only revisions of existing items were considered. Hereafter, the pilot testers attended a one-day training course, pilot tested the MISA-DK at own facility, and responded to the CVQ.

For the content identification, the two ICF-experts were introduced to the final version of MISA-DK and followed established linking rules: a) identification of all meaningful concepts within the overall purpose and the items and score descriptors of the MISA-DK and b) linking of each meaningful concept to the most precise ICF category (61,62,89).

In Study II, the MISA- DK was administered to the patients at breakfast or lunch time as in-person observation by TH. In addition, the patients' performance during the meal was video-recorded for the reliability study. Hereafter and within 2 days, data for the convergent and known-groups validation were collected by a research assistant (an experienced occupational therapist). At the time of

the data collection, the research assistant was blinded to the results on the MISA-DK.

Convergent validity was determined by 4 constructs to cover the complexity of ingestion (35,41):

1) Cognition measured with the Mini-Mental Status Examination (MMSE) with a score range of 0 to 30. Decreasing scores indicate reduced cognitive function (90,91).
2) Physical function measured with the Barthel-100 index (BI) with a score range of 0-100. Decreasing scores indicate reduced physical function (92).
3) Orofacial function measured with the Nordic Orofacial Test-Screening (NOT-S) with a score range of 0 to 6 for the clinical examination. Higher score indicates orofacial dysfunction (93).
4) Swallowing function measured with a passed or failed Water swallow test (WST) (94).

Known-groups validity was examined in terms of frailty and pulmonary status. The patients were considered frail if they fulfilled three or more of five criteria (4):

1) Weight loss: determined by the initial screening of the Nutrition Risk Screening (95) routinely performed and documented by the facilities' nursing staff.
2) Exhaustion: determined by a score <50 at the Danish version of the WHO-five Well-Being index (WHO-5). The score ranges from 0 to 100 (96).
3) Weakness: determined by decreased grip strength measured with a handheld dynamometer (average of 3 measures using dominant hand) and established norms at age and gender (97).
4) Slowness: determined by a time of >19 seconds on the "Timed Up & Go" test (98).
5) Poor physical activity: determined by a BI score <50, indicating moderate to severe functional disability (92).

The pulmonary status in terms of pneumonia was determined on basis of the diagnosis made by the medical physician of the patient and documented in the patients' medical file.

In Study III, the video-recordings from the mealtime observations in study II were used for the rater reliability study. In this way, independence between raters as well as the stability in the patients' performance was ensured (36,42-45,66). The video-recordings were saved into a CD in mpeg format and lasted on average 24 minutes (range: 8 to 43 minutes). In order to minimize rater-errors, all raters underwent a training course before participation in the study (99-101). Details on the course are given in paper III. The raters were paired randomly across the clinical settings in a two-rater design for each video-case (interrater reliability) (36,42-45), and scored on average 5 video-cases (range 2 -11). The rater re-scored the same video-cases in a test-retest design within a time frame of

3 to 8 weeks (intrarater reliability) (36,42-44). Each rater received continuously the CD's and MISA-DK score-sheets with information on basic patient demographics and diagnostics and of the mealtime menu.

4.4. Analysis

Statistical analyses were done by means of SAS 9.1 (Strategic Analysis System), SPSS 17.0 and 19.0 (Statistical Package of Social Science) and RUMM2030 (Rasch Unidimensional Measurement Models) (102). Descriptive analysis of demographics and clinical measures were based on frequencies, mean, range and SD (Study I-III). The level of statistical significance was 5% and was two-sided for all comparisons among groups.

Study I

The responses to the Content validity questionnaire in the fourth phase of the translation and in the pilot test were analysed using the Content validity index (CVI) (88,103) and the Average deviation (AD) index (104) (see Paper I for definitions). Adequate content validity of MISA-DK required a CVI of 1.0 (105) and an AD index < 0.65 for the expert panel and < 0.67 for the pilot testers (104). Comparisons of the judgments made by the pilot testers who also participated as content experts versus those who did not were analysed using the Mann–Whitney U-test (44,106). For the content identification, the number of meaningful concepts, the linked ICF categories and their distribution within the ICF components were calculated. Content density was estimated as the ratio of the number of identified concepts and the number of items; a value > 1 may indicate complex items. Content diversity was estimated as the ratio of the number of linked ICF categories and the number of identified concepts; a value < 1 may indicate that several concepts and their items are dedicated to the same topic (89).

Study II

This summary contains an extension of the undertaken Rasch analysis in study II, which resulted in supplementary analyses within CTT on data from study II and III. In order to maintain a logical sequence throughout the summary, the analyses of the criterion-related construct validity by means of Rasch analyses are presented first.

Criterion-related construct validity, Study II: A Likelihood-ratio test (102) on data from all 43 MISA-DK items revealed that the partial credit model (81) should be used. The initial Rasch analysis in study II treated all 43 items as a total scale, which were analysed via a multistep process (64,

73,79,107). Overall model fit was considered with: a) summary fit residual statistics for items and persons, which should have a mean close to 0.0 and a SD of 1.0 (SD < 1.4 is usually accepted); and b) summary of the item chi-square (χ^2) statistics, which should be non-significant ($p > 0.05$) reflecting the invariance of the items across different ability groups (64,107). Reliability and power of fit was considered using the Person-Separation Index (PSI) (107). A PSI >0.7 is required (108). Unidimensionality was analysed using t-tests to compare person ability estimates derived from the two most disparate subsets of scale items, which were created from principal component analysis of the residuals. Unidimensionality is supported if less than 5% of cases show a significant difference or if the value of 5% falls within the 95% CI (107,109,110).

Sources of deviation from model expectation were examined to see if the MISA-DK could be improved. Thresholds ordering were considered using the threshold map and category probability curves (64,107). Disordered thresholds were resolved by merging adjacent categories (79,82,83). Individual item and person fit was considered using fit residuals in the range of ±2.5 and/or χ^2 and F-statistical probability values. Misfitting items were removed to try to improve overall model fit. If not, misfitting items were retained (64,107). Local independency was investigated using the residual correlation matrix of the items (102). Local item dependency (LID) was evident by item residual correlations above 0.2 (111), and was dealt with by grouping local dependent items into a testlet (a higher-order item), which absorbs the impact of LID (111-113). DIF analysis was undertaken for the person factors gender (male, female) and age (defined by the median of 83 years). DIF was analysed via a 2-way analysis of variance (ANOVA) on the residuals for each item/testlet across the person factors and across the class intervals testing the main effect (uniform DIF) or an interaction effect (non-uniform DIF) (64,102,114). Uniform DIF reflects significant item difficulties between groups, and can be adjusted by splitting the DIF item into group specific items (64,107,114). Non-uniform DIF is usually removed as it reflects significant difference in item discrimination between groups, i.e. misfit to the model across the continuum (64,107).
The Bonferroni correction was used to adjust for multiple testing (overall and individual item fit and DIF), keeping the Type I error to 5% (44,64,106,107)
When the data fitted the model, the scale to sample targeting was evaluated using the Person-item-threshold distribution (64,73,107). A sample size of 100 patients who are reasonable well targeted will provide 95% confidence that the estimated item difficulty is within ±0.5 logits (115).

Supplementary analysis: In study II, the Rasch analyses revealed that considerable LID was identi-

fied for items within each subscale, which was accommodated through the testlet design (111-113). To verify whether the items within each testlet fitted the Rasch model, this summary presents an extended Rasch analysis of the individual subscales. Except for items presenting non-uniform DIF, which was initially removed from the scale, the fitting solutions included adjustments of LID and stepwise removal of misfitting items presenting the greatest magnitude of misfit.

Convergent validity, Study II: Spearman's rho (r_s) was used as neither of the variables were normal distributed (44,106). Adequate correlations ($r_s > 0.50$) (44) were expected for: a) the MISA-DK total scale and all 4 convergent variables; b) the Positioning subscale and the BI; c) the Self-feeding skills subscale and the BI and the MMSE; and d) the Solid and Liquid ingestion as well as the Texture management subscales and the NOT-S and the WST, respectively. In addition, stepwise multiple regression analysis was applied to assess the relative importance and contributions of the convergent variables to variance in ingestive skills ability (63).
Supplementary analysis: The extended Rasch analysis on the MISA-DK subscales revealed that supplementary analyses of the convergent validity on some single items were needed. Thus, rank-biserial correlations for binary and ordinal measures (44) were applied in relation to the WST.

Known-groups validity, Study II: The Mann Whitney U-Test was used for the MISA-DK subscale and total scores, and for ordinal scores (supplementary analyses) (44,106).

Internal consistency reliability, Study II: Cronbach's α was calculated for the MISA-DK items within each subscale and for the total scale. Values of 0.70 to 0.90 are acceptable (43,44,63). In this summary, Cronbach's α is presented together with the reliability estimates in the Rasch analyses.

Study III

Inter- and intra-rater reliability, Study III: Relative reliability was estimated for the MISA-DK subscales and total score using the ICC; model 1 ($ICC_{1.1}$) for interrater, and model 3 ($ICC_{3.1}$) for intrarater reliability (44,65). For the $ICC_{1.1}$, the consistency definition was applied because the variance due to systematic differences between raters is partitioned out in its calculation (65). For the $ICC_{3.1}$, the absolute agreement definition, which includes the rater variance, was applied (65). ICC values >0 .75 indicate excellent reliability and ICC values between 0.60 and 0.74 indicate good reliability (116). A sample size of 102 patients was estimated to obtain ICC > 0.75 with a lower 95% CI > 0.60. A power of 80% and α of 0.05 were used (117).

Absolute reliability was estimated using the standard error of measurement (SEM) and the smallest detectable change (SDC) (43-45,65,67,68). The SEM was calculated from the ANOVA statistics when computing the ICC (65,67,69) and was considered small if it represents ≤10% of the absolute scale range (118). The SDC was calculated from the SEM (65,67,68). To examine whether the error of measurement was dependent on the magnitude of the mean score (heteroscedasiticity), Bland-Altman plots for the rater-pairs and for the two time points were constructed (119-121), and Limits of agreement (LOA) were calculated (119).

Supplementary analysis: The extended Rasch analysis on the MISA-DK subscales revealed that a supplementary analysis of the rater reliabilities on some single items was needed. As the $ICC_{1.1}$ for five of the six subscales did not exceed 0.75, all items in the MISA-DK were analysed in terms of inter-and intrarater reliability. Percentage of observed agreement (PO) and quadratic weighted Kappa (Kw) was calculated (66,116,122,123). PO ranges from 0 to 100; PO < 70 is considered poor reliability; 70-79 is fair; 80-89 is good and 90-100 is excellent (116,124). The calculation of Kappa is based on the difference between the observed agreement compared to how much agreement would be expected to be present by chance alone (43,44,116,122,123). Kappa ranges from -1 to 1; Kappa <0.4 is considered poor reliability, 0.40-0.59 is fair, 0.61-0.74 is good, and more than 0.75 is excellent (116).

4.5. Ethical considerations

The study was approved by the Danish Data Protection Authority (Reg. No: 2009-41-3719) and the local Scientific Ethical Committee in the Capital region (Reg. No: H-C-2009-061), and was registered in the Clinical Trial Database (Reg. No: NCT01006330). All participants gave written informed consent regarding participation.

The CDs with the video-recordings were treated confidentially, and were stored in a locked cabinet when not in use. When in use, the patients were made anonymous on the corresponding MISA-DK score-sheets. The CDs and the MISA-DK score sheets were personally delivered to the rater by TH. In addition, all raters signed a sworn statement regarding maintaining confidentiality and that the videos would not be accessible to unauthorized persons. The methods applied in the study were considered not to give rise to any ethical problems in regard to the status of autonomy, integrity or physical well-being of the participants. The participation of the patients was scheduled allowing breaks in the assessment process. If any procedure was deemed to be dangerous to the patient or if the patient did not want to continue, it was terminated.

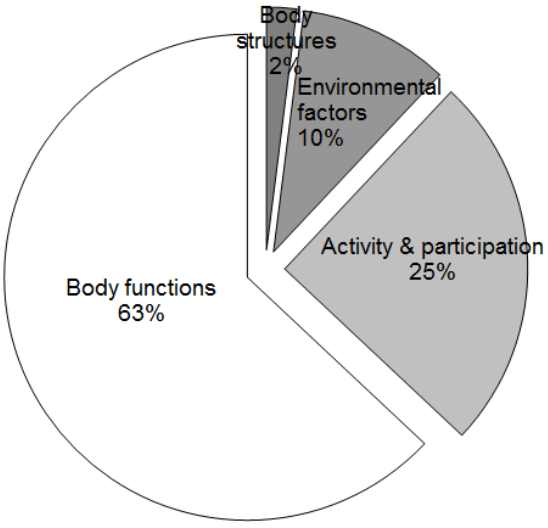
5. Results

5.1. Translation and cultural adaptation of the MISA into MISA-DK. Study I

The first three phases of the translation resulted in a MISA-DK with few cultural adaptations adressing some of the solid consistencies in the texture-management subscale. Phase four revealed: 40 adequate item terms; 29 clear item descriptors; 37 clear score descriptors and 40 relevant items when using CVI of 1.0 and the AD < 0.65 as cut points for adequate content validity. The not content valid domains involved predominantly the clarity of the texture-management items and their scoring. All sections in the instruction manual and the score sheet, but one (the equations for calculation of a percentage score) was judged to be clear. Based on the analyses and discussions with the content-expert and Heather C. Lambert, some modifications were undertaken (See the translation report in Appendix A). The subsequent pilot test revealed that the MISA-DK, but one item in the positioning subscale ("Maintain 90-degree hip angle") appeared content valid when using CVI of 1.0 and AD < 0.67 as cut points; no comments were made and no further modifications were undertaken. This resulted in a final version of the MISA-DK to be psychometrical tested (Appendix B). There were no significant differences in the judgments of the content validity domains between the pilot testers who also participated as content-experts versus those who did not (Range, p=0.06 to 1.0; Interquartile range, p=0.29 to 1.0).

Content identification and quantification: The overall purpose of the MISA-DK was linked to the ICF categories eating (d550) and drinking (d560) within the activity and participation component. In total, 214 meaningful concepts were identified (177 belonged to the texture management subscales) and linked to 41 ICF categories. Their distribution is given in Figure 1. The addressed categories were related to: ingestion-, mental-, sensory-, voice and speech -, respiration -, neuromusculoskeletal and movement-related functions, learning and applying knowledge, mobility, interpersonal interactions and relationships, products and technology, and support and relationship. The density ratio was 5 and the diversity ratio was 0.2.

Figure 1. Distribution of 41 ICF categories



5.2. Psychometrical properties of the MISA-DK. Study II & III

Criterion-related construct validity, Study II: The main results of the Rasch analyses are provided in Table II at the next page. Initially, the original MISA-DK did not fit to the Rasch model (Table II; Analysis 1). Removal of six items showing misfit by one or more of the three fit statistics did not produce model fit and they were retained. Model fit was achieved after resolving disordered thresholds for 11 items by merging score 1 and 2, and by creating six testlets of LID items with correlated residuals clustering within each subscale (Table II; Analysis and 2). No DIF by gender or age was presented at item or testlet level.

The extended Rasch analysis on the individual MISA-DK subscales and the different fitting solutions are shown in Table II (Analysis 3 to 13). The Positioning subscale was initially consistent with Rasch model expectations (Table II; Analysis 3). For the Self-feeding skills subscale, item 9 manifested non-uniform DIF by gender and was removed (Table II; Analysis 4 and 5). The Liquid ingestion subscale showed reasonable fit to the model, but improved further after LID was adjusted by two testlets. This however, decreased the reliability markedly (Table II; Analysis 6 and 7). The Solid ingestion subscale showed reasonable fit to the model, but improved further after LID was adjusted by four testlets (Table II; Analysis 8 and 9). The Texture management-solids subscale showed misfit to the Rasch mode (Table II; Analysis10). Adjustment of LID by one testlet worsened the overall item fit statistics (Table II; Analysis 11). Stepwise removal of misfitting items, starting with the testlet of items 31/33 having the poorest fit, improved model fit, but with too low power and a high number of extreme scores (Table II; Analysis 12). The Texture management liquid subscale did fit the Rasch model, but the power of fit was too low and the number of extreme scores was high (Table II; Analysis 13). For the six subscales, the mean location of persons ranged from 0.693 to 1.541, and the majority of the extreme scores were at the ceiling.

The consequences of Analysis 3 to 13 are that item 9 and the texture management items are regarded as single items (item location and fit residuals are given in Appendix C). In order to investigate, whether a total score could validly be obtained by the remaining 29 items, they were re-analysed as a total scale. This revealed reasonable model fit, but indication of multidimensionality. Item 16 presented marginally non-significant misfit. It was retained as removal worsened the overall model fit. Adjusting LID among clusters of items within each subscale improved overall model fit and the unidimensionality t-test (Table II; Analysis 14 and 15). Thus, supplementary analyses within CTT was needed on the Rasch revised MISA-DK total scale, the Self-feeding skills subscale, item 9 and the single items within the texture management scales.

Table II Rasch model fit statistics of the 43 MISA-DK items as one scale and as distributed into the original six subscales

Analysis		Item fit residual	Person fit residual	Chi-square interaction		PSI	α	T-tests in % (95%CI)	n
		Mean (SD)	Mean (SD)	Value (df)	*P*				
MISA-DK scale (43 items)									
1	Misfit of 6 items; DT of items 7,12,13,32,34,35,37,38,39,42, 43; LID of items within sub-scales; no DIF.	-0. 02 (1.81)	0.10 (1.18)	312.5 (86)	<0.001	0.93	0.95	21.8 (17.7;25.9)	110
2	Testlets of 6 subscales after rescoring of 11 Items → No item misfit; no DIF; no LID.	-0.29 (1.35)[§]	-0.35 (0.94)	12.3 (12)	0.424	0.85	0.88	4.6	110
Positioning (4 items)[‡]									
3	No item misfit; no DIF; no LID; no DT.	-0.41 (0.63)	-0.37 (0.72)	9.8 (8)	0.281	0.71	0.79	N/A*	92
Self-feeding skills (7 items)[‡]									
4	LID of items 5/6; non-uniform DIF (gender) of item 9; DT of item 7[¤].	-0.40 (1.59)	-0.31 (0.91)	33.8 (14)	0.002	0.71	0.85	2.0	98
5	Remove item 9-Non-uniform DIF → No item misfit, no LID; DT of item 7[¤].	-0.31 (1.41)	-0.39 (0.99)	20.7 (12)	0.054	0.72	0.86	3.2	93
Liquid ingestion (7 items)[‡]									
6	Non-significant misfit of item 16 (FR=2.9); LID of items 12/14/15 and 17/18; DT of item 12[¤].	-0.33 (1.49)	-0.36 (0.75)	16.1 (14)	0.310	0.73	0.83	3.3	91
7	Two testlets of items with LID → No DIF or item misfit.	0.26 (1.22)	-0.23 (0.66)	7.1 (10)	0.521	0.48	0.70	2.2	91
Solid ingestion (12 items)[‡]									
8	No item misfit; LID of items 21/23, 24/25, 26/27, 29/30; uniform DIF (gender) of item 29; no DT.	-0.11(1.22)	-0.15 (1.04)	39.5 (24)	0.024	0.86	0.90	4.9 (0.7;9.2)	100
9	Four testlets of items with LID → No DIF or item misfit.	0.69 (1.14)	0.04 (0.89)	27.1 (16)	0.041	0.82	0.84	5.0 (0.7;9.3)	100
Texture management-solids (8 items)[‡]									
10	Misfit of items 31,33,37,38; LID of item 31/33; DT of items 37,38[¤]; no DIF.	0.09 (1.98)	-0.01 (1.36)	121.5 (16)	<0.001	0.72	0.80	4.1	98
11	One testlet of items with LID → Misfit of testlet 31/33 and items 32, 35,36,37,38.	0.06 (2.10)	-0.06 (1.26)	85.3 (14)	<0.001	0.65	0.75	1.0	98
12	Stepwise removal of misfitting items → No DIF; no misfit of remaining items 32.35,36.	0.27 (1.53)	-0.33 (1.65)	2.6 (6)	0.862	0.28	0.70	N/A*	60
Texture management-liquids (5 items)[‡]									
13	Misfit of item 41; LID of items 39/40, 42/43; no DIF; DT of item 43[¤].	0.91 (0.56)	-0.21 (1.75)	18.8 (10)	0.043	0.28	0.77	N/A*	59
Rasch revised MISA-DK total scale (29 items)[‡]									
14	LID of items within subscales; no DIF; Non-significant misfit of item 16 (FR=3.2); DT of items 7,12,13[¤].	-0.12 (1.21)	-0.17 (1.10)	90.5 (58)	0.004	0.92	0.95	20.6 (16.4;24.7)	107
15	Testlets of 4 subscales→ No item misfit; no DIF; no LID.	-0.54 (1.88)[§]	-0.46 (1.01)	13.0 (8)	0.112	0.82	0.83	1.87	107
Ideal fit		**0.0 (1.4)**	**0.0 (1.4)**		**>0.05[#]**	**>0.7**	**>0.7**	**<5.0**	

Abbreviations: PSI, Person separation index; α, Cronbach's alpha; n, numbers of patients without extreme scores included in the analysis; DT, Disordered thresholds; LID, Local item dependency; DIF, Differential item function; FR, Fit residual .
Notes: [‡]Extended analysis for this summary; [¤]Not rescored as it worsened model fit; [#]Bonferroni adjusted; *N/A: Not applicable because of too few items in one of the subsets; [§] The item residual SD may be inflated when testlets are of different length (128).

Convergent validity, Study II: The correlations of the MISA-DK and the convergent variables before and after the extended Rasch analysis are displayed in Table III.

Table III. Correlations of the original and Rasch revised MISA-DK and the convergent variables

MISA-DK scales (Number of items)	MMSE	BI	WST	NOT-S
Positioning subscale (4)[#]	0.43**	**0.70****	0.10	-0.34**
Self-feeding Skills subscale (7) [#]	**0.54****	**0.66****	0.15	-0.45**
(Rasch revised Self-feeding skills subscale (6))[‡]	**0.52****	**0.66****	0.22*	-0.50**
Liquid ingestion subscale(7) [#]	0.42**	0.40**	0.39**	-0.39**
Solid ingestion subscale (12) [#]	**0.51****	**0.61****	0.31**	**-0.52****
Texture management-solids subscale (8) [#]	**0.50****	0.41**	0.24*	-0.39**
Texture management-liquids subscale (5) [#]	0.41**	0.32**	0.21*	-0.25*
MISA-DK total scale (43) [#]	**0.59****	**0.66****	0.31**	**-0.53****
(Rasch revised MISA-DK total scale (29))[‡]	**0.54****	**0.68****	0.31**	**-0.52****
Single items in the Rasch revised MISA-DK[‡]				
Able to focus on meal	0.35**	0.32**	-0.08	-0.03
Capable of eating heterogeneous textures	**0.57****	**0.50****	0.22**	-0.49**
Capable of eating fibrous solids	0.45**	0.41**	0.12	-0,39**
Capable of eating hard solids	**0.50****	0.44**	0.28**	-0.44**
Capable of eating minced/granular solids	0.26**	0.35**	0.11	-0.24*
Capable of eating sticky solids	0.46**	0.30**	0.18	-0.26**
Capable of eating soft	0.46**	0.34**	0.24**	-0.34**
Capable of eating puree	0.17	0.18	0.08	-0.08
Capable of eating pudding	0.09	0.15	0.10	-0.06
Capable of drinking water	**0.52****	0.40**	0.34**	-0.32
Capable of drinking thin juices	0.47**	0.44**	0.26**	-0.37**
Capable of drinking nectar consistency	0.43**	0.38**	0.21*	-0.37**
Capable of drinking honey consistency	0.12	0.11	0.07	-0.04
Capable of drinking pudding	0.23*	0.11	0.15	-0.08

Shades= hypothesized adequate correlations, bold =adequate correlations by Spearman's rho or rank-biserial correlations > 0.50.
** p<0.01; *p<0.05.
[#]Results are obtained before the Rasch analyses and are published in paper II.
[‡]Results based on supplementary analyses after the extended Rasch analyses presented in this summary.

The multivariate analysis of the original MISA-DK revealed that of the six subscales, the total explained variance was least for liquid ingestion (37%), texture management solids (35%) and texture management liquids (21%).

Known-groups validity, Study II: When measured with the original MISA-DK, frail patients (n=64) obtained significantly lower scores in all six subscales and the total score than robust patients (n=40) (p<0.001). A statistically significant difference persisted for the Rasch revised MISA-DK total scale (Z=-5.7, p<0.001) and the Self-feeding skills subscale (Z=-5.31, p<0.001), as well as for the single item 9 (Z=-2.60, p=0.009). For the single texture management items, a difference was found for nine items (Range, Z=-4.28 to Z =-2.23; p= <0.001 to p= 0.026). No difference was found for item 37, 38, 42, 43 (Range, Z=-1.06 to Z=-1.69; p=0.092 to p=0.289).

When measured with the original MISA-DK, patients with pneumonia (n=48) obtained significantly lower scores in liquid ingestion (p=0.029), solid ingestion (p=0.047), texture management solids

(p=0.007) and the total scale (p=0.034) than patients without a diagnosis of pneumonia (n=62). For the Rasch revised MISA-DK in terms of the single texture management items, a statistically significant difference was found for eight items (Range, Z=-2.04 to Z =-2.63; p=0.008 to p=0.041). No difference was found for item 36,37,38,42,43 (Range, Z = -0.13 to Z =-0.762; p=0.138 to p=0.894) and the Rasch revised MISA-DK total scale (Z=-1.95, p=0.051). No difference persisted for the Rasch revised Self-feeding skills subscale (Z=-1.67, p=0.096).

Inter- and intrarater reliability, Study III: The MISA-DK demonstrated good to excellent relative interrater reliability and excellent intrarater reliability before and after the extended Rasch analysis (Table IV).

Table IV Inter- and intrarater reliability of the original and Rasch revised MISA-DK

	Interrater reliability				Intrarater reliability			
	Relative reliability		Absolute reliability		Relative reliability		Absolute reliability	
MISA-DK scales (scale range)	$ICC_{1.1}$	95% CI	SEM%	SDC	$ICC_{3.1}$	95% CI	SEM%	SDC
Positioning subscale (4-12)[#]	0.61	0.47-0.72	15%	3.3	0.87	0.83-0.90	9%	1.9
Self-feeding skills subscale (7-21)[#]	0.71	0.60-0.80	11%	4.4	0.86	0.82-0.89	8%	3.0
(Rasch revised Self-feeding skills subscale(6-18))[‡]	*0.76*	*0.67-0.83*	*11%*	3.6	*0.88*	*0.84-0.91*	*8%*	2.6
Liquid ingestion subscale (7-21)[#]	0.73	0.63-0.81	9%	3.6	0.89	0.85-0.91	6%	2.5
Solid ingestion subscale (12-36)[#]	0.73	0.62-0.81	10%	6.7	0.88	0.84-0.91	7%	4.4
Texture management-solids subscale (8-24)[#]	0.74	0.63-0.81	14%	6.1	0.84	0.80-0.88	10%	4.4
Texture management-liquids subscale (5-15)[#]	0.76	0.66-0.83	14%	3.9	0.88	0.85-0.91	10%	2.8
MISA-DK total scale (43-129)[#]	0.84	0.77-0.89	7%	15.8	0.93	0.90-0.94	4%	10.3
(Rasch revised MISA-DK total scale (29-87))[‡]	*0.80*	*0.71-0.86*	*8%*	*12.2*	*0.92*	*0.90-0.94*	*5%*	*7.5*

SEM%; standard error of measurement as a percentage of the absolute scale range; SDC; smallest detectable change
All ICC were significant $p < 0.001$
[#]Results are obtained before the Rasch analysis and are published in paper III.
[‡] Results based on supplementary analyses after the extended Rasch analyses presented in this summary.

For the absolute reliability, the SEM% and SDC were larger between than within raters (Table IV), and hence the LOA's were broader between raters (Paper III, Table II and III). The Bland-Altman plots did not indicate heteroscedasiticity (Paper III, Figure 1).

The results of the supplementary reliability analysis at item level are provided in Table V at the next page. Adequate PO values above 70 were present for 20 items between raters and 43 items within raters. Good to excellent PO values above 80 were present for seven items between raters and 30 items within raters. Adequate point estimates of Kw above 0.4 were present for 36 items between raters and 43 items within raters. Good to excellent point estimates of Kw above 0.60 were present for 15 items between raters and 41 items within raters.

Table V. Supplementary analysis of inter- and intrarater reliability of each MISA-DK item by Percentage of observed agreement (PO) and weighted Kappa (Kw).

	Interrater		Intrarater	
Positioning	PO	Kw (95%CI)	PO	Kw (95%CI)
1. Maintain symmetry of posture	70	0.50 (0.31;0.62)	84	0.77 (0.67;0.86)
2. Maintain adequate head positioning for feeding	76	0.61 (0.42;0.76)	87	0.75 (0.63;0.86)
3. Maintain 90-degree hip angle	77	0.14 (-0.04;0.33)	92	0.77 (0.63;0.83)
4. Able to sit upright without leaning on arm	56	0.45 (0.29;0.62)	75	0.75 (0.63;0.83)
Self-feeding skills				
5. Able to grasp utensil functionally and bring it to the mouth	77	0.74 (0.63;0.85)	82	0.76 (0.67;0.85)
6. Able to grasp cup/glass functionally and bring it to the mouth	90	0.79 (0.62;0.95)	95	0.91 (0.84;0.98)
7. Selects appropriate utensil for food item	93	0.81 (0.64;0.98)	92	0.80 (0.68;0.92)
8. Takes appropriately-sized mouthfuls	50	0.32 (0.13;0.51)	75	0.69 (0.60;0.78)
9. Able to focus on meal	56	0.41 (0.26;0.56)	74	0.65 (0.54;0.75)
10. Demonstrates good judgment	50	0.34 (0.19;0.50)	74	0.68 (0.60;0.76)
11.Tolerates physical effort of meal	60	0.59 (0.48;0.70)	83	0.79 (0.71;0.87)
Liquid ingestion				
12. Seals lips on cup/glass	87	0.46 (0.13;0.79)	95	0.79 (0.64;0.94)
13. Able to draw liquid from a standard straw	88	0.82 (0.70;0.95)	93	0.92 (0.87;0.96)
14. Prevents leakage of liquid from cup/glass while drinking	85	0.54 (0.25;0.83)	89	0.66 (0.49;0.84)
15. Prevents leakage of liquid from mouth before swallowing	86	0.34 (0.15;0.52)	89	0.59 (0.43;0.76)
16. Able to take a sequence of sips	51	0.42 (0.26;0.58)	81	0.80 (0.73;0.88)
17. Demonstrates same voice quality after drinking	59	0.51 (0.37;0.66)	78	0.72 (0.63;0.82)
18. Clear airway if necessary after liquids	62	0.49 (0.32;0.66)	80	0.73 (0.64;0.83)
Solid ingestion				
19. Close upper lip on utensil	73	0.45 (0.23;0.66)	82	0.67 (0.56;0.79)
20. Prevents the loss of food from the mouth before swallowing	68	0.43 (0.28;0.58)	82	0.64 (0.52;0.75)
21. Use functional chewing pattern	66	0.37 (0.15;0.59)	90	0.82 (0.73;0.91)
22. Chewing appropriate to food item	68	0.43 (0.23;0.63)	81	0.72 (0.64;0.81)
23. Position bolus when chewing	69	0.46 (0.25;0.67)	87	0.82 (0.74;0.89)
24. Quantity of food remaining in mouth after swallow	70	0.53 (0.35;0.70)	81	0.70 (0.60;0.81)
25. Location of food remaining in the mouth after swallow	70	0.54 (0.38;0.71)	84	0.73 (0.62;0.84)
26. Swallow without extra effort	62	0.43 (0.26;0.60)	70	0.60 (0.49;0.71)
27. Swallows only once or twice per mouthful	59	0.20(-0.01;0.40)	77	0.62 (0.52;0.73)
28. Maintain respiratory pattern throughout meal	64	0.60 (0.46;0.74)	79	0.82 (0.76;0.87)
29 .Demonstrates same voice quality after eating	64	0.55 (0.39;0.70)	79	0.70 (0.60;0.80)
30. Clear airway if necessary after solids	65	0.50 (0.33;0.67)	78	0.73 (0.63;0.82)
Texture management-solids				
31. Capable of eating heterogeneous textures	76	0.75 (0.64;0.86)	76	0.77 (0.68;0.86)
32. Capable of eating fibrous solids	68	0.64 (0.50;0.78)	82	0.80 (0.72;0.88)
33. Capable of eating hard solids	65	0.61 (0.47;0.76)	83	0.73 (0.63;0.83)
34. Capable of eating minced/granular solids	63	0.51 (0.34;0.67)	80	0.63 (0.52;0.74)
35. Capable of eating sticky solids	71	0.63 (0.48;0.78)	77	0.74 (0.65;0.83)
36. Capable of eating soft solids	66	0.35 (0.12;0.57)	81	0.52 (0.37;0.66)
37. Capable of eating puree	66	0.48 (0.31;0.64)	77	0.71 (0.62;0.80)
38. Capable of eating pudding	64	0.42 (0.25;0.59)	81	0.78 (0.70;0.87)
Texture management-liquids				
39. Capable of drinking water	74	0.73 (0.61;0.86)	85	0.85 (0.79;0.92)
40. Capable of drinking thin juices	76	0.73 (0.61;0.86)	87	0.83 (0.76;0.91)
41. Capable of drinking nectar consistency	75	0.63 (0.48;0.80)	86	0.81 (0.72;0.90)
42. Capable of drinking honey consistency	81	0.79 (0.68;0.91)	87	0.88 (0.83;0.94)
43. Capable of drinking pudding consistency	79	0.73 (0.60;0.87)	89	0.84 (0.77;0.91)

Kw; weighted Kappa using quadratic weights (71,121,127)
Reference values for PO: <70=poor; 70-79=Fair; 80-89=good; 90-100=Excellent (124)
Reference values for Kw: <0.40=poor; 0.40-0.59=fair; 0.60-0.74=good; 0.75-1.0=excellent (116).

6. Discussion

6.1. Main findings of the three studies

The studies in this thesis aimed to generate a functional equivalent Danish version of MISA, which possesses adequate levels of validity and reliability. In study I, a content valid MISA-DK was produced through a collaborative translation approach, expert-panel judgment and pilot-testing. Additional information on the content was provided using the ICF as a frame of reference, and it was found that the content density was high and the content diversity was low. In study II, Rasch analysis revealed that the MISA-DK initially did not measure a unidimensional construct. When adjusting disordered thresholds and LID throughout the scale, fit to Rasch model was achieved. During the extended Rasch analysis, it was possible to achieve model fit for four of six subscales. The two Texture management subscales did not succeed adequate model fit, and their items were considered as single items. A Rasch revised MISA-DK total scale achieved model fit after adjusting LID throughout the scale. When using analyses within CTT before and after the Rasch analyses, the results provided support for the internal consistency reliability. The convergent validity was supported for the Positioning- and Self-feeding skills subscales and partially supported for the Solid ingestion subscale and the MISA-DK total scale. It was not possible to establish convergent validity in terms of orofacial and swallowing functions of the Liquid ingestion subscale and the texture management items. The known-groups validity was supported for the MISA-DK total scale, subscales and some of the texture management items. In study III, the MISA-DK total scale and subscales exhibited good to excellent interrater reliability and excellent intrarater reliability. The amount of measurement error was small for the MISA-DK total scale, but relatively large for the subscales between raters. The supplementary item level reliability analysis of MISA-DK revealed that some items demonstrated poor interrater reliability. A contribution of these studies is accumulated validity evidence on the MISA-DK. Brown (54), Smith (55) and Wilson (56) outlines the used validation activities within CTT and the Rasch model for the unified concept of construct validity; and the study results are discussed within this frame.

Test content

Prior to the study, a literature review was undertaken (25), and MISA was found to include relevant and representative items for measuring occupational performance in eating and drinking as defined within Danish occupational therapy. In study I, the production of a functional equivalent translation of MISA was initiated using a collaborative translation approach in many steps (57). Such alternative approaches have shown to be as good as the back-translation approach for patient reported out-

come measures (PROMs) (125,126). Linking the MISA-DK to the ICF revealed that the content reflects relevant aspects of the construct "ingestion". However, potential problems might be inherent in the 13 texture management items, as the expert panel suggested that several assessment purposes were present at the same time in terms of swallowing safety and patient willingness (Appendix A), and a high number of meaningful concepts were identified. Thus, the content of these items may be ambiguous (89), and might have influenced on the high content density estimate of 5 across all items in the MISA-DK.

In study II, evidence of whether the content of the MISA-DK total scale is an adequate and representative reflection of the measured construct (54-56) could be achieved after adjustment of LID during the initial and extended Rasch analyses. The PSI was kept above 0.80 implying that three ability levels were identified, which is sufficient for interpreting the construct defined by the testlets (55,108). The extended Rasch analyses of the individual subscales revealed that a sufficient number of ability levels were identified for the Positioning-, Self-feeding skills- and Solid ingestion subscales, but not for the Liquid ingestion subscale when adjusting LID. This resulted in a PSI= 0.48, which is not sufficient (108). The LID items represent swallowing efficiency in terms of the ability to control the lips while drinking (i.e., item, 12,14,15) and swallowing safety in terms of the ability to protect the airway from aspiration (i.e., item 17,18). However, items representing swallowing efficiency in terms of bolus propulsion (5,28,30) are not represented in the Liquid ingestion subscale; which is the case for the Solid ingestion subscale. If the expert panel in study I. had judged the content representativeness of the items (59) in addition to the content relevancy (50), it cannot be excluded that this would have turned up, and the content validity would have been covered more adequately (54,56,59,60). Nevertheless, the extended Rasch analysis indicates that the Liquid ingestion subscale might benefit of adding items covering additional aspects of swallowing efficiency.

The initial and extended Rasch analyses revealed that item 16 (able to take a sequence of sips) demonstrated multidimensionality. The item assesses the coordination of drinking and breathing (24). Since the physiology of multiple swallows is different than during a single swallow, it is an important item (22). In the item and score description, it is stated that the patient should not be asked to take a sequence of sips; since if he avoids this, it could reflect a functional loss (24). Such a description appears unclear and allows guessing. Therefore, it is suggested, that if the patient does not present extremely poor swallow-respiratory coordination during single swallows of liquids, then he is asked to drink continuously (22) during the meal.

Although the extended Rasch analysis revealed model fit of the two Texture management subscales,

the PSI and thus the power in detecting the items not fitting the model became too low (102,107). Hence, it is difficult to state whether or not the items within these two scales shares a common underlying dimension (55,64,73,107). In fact, it could be debated, whether these items act as an additional facet beyond the item difficulty parameter (79). That is, the textures to be ingested represents different task challenges during a meal (5,22). Although, the purpose of the scales is to assess the texture management of the patients, this ability in terms of swallowing efficiency and safety are covered by the items in the Liquid- and Solid ingestion subscales. Therefore, it can be suggested that the purpose of the texture management items are revised. Hereafter, validation using more complex Rasch models such as the many-facet Rasch model (79) is needed.
Finally, the mean person locations were in general higher than the average levels measured by the MISA-DK total scale and the subscales, and gabs in the item locations within each subscale exist (Appendix C). This might indicate a need for development of more items (54-56,64,107).

Response processes

In study II, evidence of the response processes of the MISA-DK total scale was supported by adequate person fit statistics and no extreme scores (54-56) in the initial Rasch analysis and when creating testlets adjusting LID. However, the extended Rasch analyses of the subscales revealed problems with extreme scores, which for the most parts scored at the ceiling for all patients. A sample with more patients at lower levels of ingestive skills ability across the subscales may reduce the observed ceiling effect. Nevertheless, this effect does raise concerns about the targeting of the scales for elderly medical patients (54-56,64,107).
The initial and extended Rasch analyses of the MISA-DK revealed that disordered thresholds were evident for some items within the Self-feeding skills-, Liquid ingestion- and Texture management subscales. This reflects that the item response categories do not operate as intended (82,83) and might benefit from revisions in order to reflect successively more of the underlying trait they are measuring. For the texture management items, ambiguous score descriptors were emphasized in the above discussion on the test content. When assessing swallowing amongst dysphagic patients, emphasis must be on swallow efficiency (bolus propulsion) and safety (airway protection) (5,28,30), and not on willingness. It can be suggested that these items simply are rated according to whether or not the patient's swallow is efficient and safe, respectively; or as discussed in relation to the test content, their purpose in the MISA-DK are reconsidered.

Internal structure

The dimensionality aspect of the internal structure (54-56) was initially addressed in study I, where

the MISA-DK was linked to categories across four ICF components, which could indicate multidimensionality (71,72). This was confirmed in study II. It was found that the MISA-DK, at first, did not fit a unidimensional Rasch model. Fit to the model and evidence of unidimensionality was provided by the creation of six testlets in order to absorb LID among items within each subscale. No item deletion, but rescoring of the response categories of 11 items was necessary; which however alters the raw score (74). Increased attention is given the testlet design to adjust the impact of LID in subscales of health outcome measurement instruments before making decisions about item deletion (111,127-130). It is argued, that by using a testlet design, the clinical utility of a scale for rehabilitation management is retained in conjunction with the fulfilment of modern psychometric standards (111, 128). However, this may be less problematic for new scales, where psychometric evidence is still accumulating (56); which can be argued is the case for MISA-DK (and MISA). From the extended Rasch analyses and as discussed above, it can be concluded that the individual subscales are far from ideal, at least in our sample of elderly medical patients, and revisions are needed.

In general, LID was a consistent feature of the MISA-DK during the initial and extended Rasch analyses. LID can be caused by response dependency or multidimensionality, which might be difficult to distinguish (112). However, when a scale is constructed by a composition of subscales, some multidimensionality might be unavoidable (131). It is argued that different content areas within a measurement instrument may impose LID on items measuring the same content area (132). As the MISA-DK subscales reflect different content areas related to ingestion, it is highly likely that content clustering may have caused the observed LID. This might be supported by the content diversity estimate of 0.2 found in study I, which indicates a relatively narrow content bandwidth of the 43 MISA-DK items (89), and the Cronbach's $\alpha > 0.90$ found in the initial and extended Rasch analyses, which indicates content redundancy (43,44,63). The applied testlet design might be regarded as being equivalent to bi-factor models, in which each item loads on two dimensions; on a main dimension and on the dimension of the unique subscale (131). In the development of the Canadian MISA, the assignation of the items into the subscales was not confirmed (47). Therefore, in order to fully understand the dimensionality of the MISA-DK and the effects of the testlets, further investigation of its dimensionality using factor analytic methods (43,44,63,110) in conjunction with multidimensional Rasch models and Rasch testlets models (131) is needed.

The Generalizability aspect of the internal structure in terms of invariance across gender and age groups was addressed in study II, and no DIF was identified during the initial Rasch analysis. However, the extended Rasch analyses revealed that item 9 (able to focus on meal) from the Self-feeding

skills subscale presented non-uniform DIF. During the initial Rasch analysis, item 9 presented a significant fit residual > 2.5, which reflects multidimensionality. Removal of item 9 during the extended Rasch analysis, improved model fit of the subscale, and the PSI and Cronbach's α increased. This might indicate that item 9 is "poor" (43,44,63). Since directed attention is a significant aspect of the ingestion construct (14,35,38,41), item 9 might benefit from revisions. Uniform DIF by gender was present for item 29 in the Solid ingestion subscale. However, this was cancelled out when the item was combined with item 30 to which LID was present. Split of item 29 by gender worsened the overall model fit, which indicates that it might not have been true DIF (114).

The reproducibility aspect of the internal structure of the MISA-DK was addressed in study III and during the supplementary analyses using CTT. The relative inter- and intra-rater reliabilities of the MISA-DK subscales and total scale were found good to excellent. However, slightly smaller magnitudes of the $ICC_{1.1}$ estimates were evident for three of the MISA-DK subscales than in Lambert et al. (48), which might be due to the sample dependency of ICC (43-45,63,65). Nevertheless, the absolute reliability estimates, which are population independent (45,65), were larger when MISA-DK was repeated by different raters than by the same rater. However, since the extended Rasch analyses revealed that a summation of the texture management items could not be justified, the reliability analyses based on their subscale scores in study III might be questionable (70-72). Yet, the supplementary item level analyses revealed relatively weaker inter- than intrarater reliability, and thus greater variation between raters than within raters. As observation based ratings are a highly complicated task, it is recognized that differences between raters' interpretation and severity will always exist (42,71,79,99,101,133). Adjusting rater severity can be realized using a many-facet Rasch model (71,79,133). A possible reason influencing our results could be contextual factors (101,133), such as different quality of the raters' computer monitors making the features of the ingestive skill items difficult to observe compared to in-person observations. In addition, unclear operational definitions of the items might provide different interpretations amongst the raters (133). This might be resolved by very comprehensive training or by modifications of the scoring instructions so they are clear and easy to use for every therapist (99,101, 133). Two of the MISA-DK items demonstrating poor interrater reliability with Kw below 0.40, namely item 3 (maintain 90-degree hip) and item 8 (takes appropriate-sized mouthfuls) were also judged unclear by the expert panel (Appendix A). Although, the PO for item 3 was fair, this might suggest a need for revisions, at least for these two items. Removing item 9 presenting non-uniform DIF, a poor PO value and a fair Kw value increased the $ICC_{1.1}$ to excellent for the Self-feeding skills subscale. This might further support a need

for revision of item 9. For the texture management items it appeared that the liquid texture items, in general, obtained stronger inter- and intrarater reliability estimates than the solid texture items. It might be due to the fact that different liquid textures are easier to categorise than solid textures (134). This may further support the aforesaid need for revisions of the texture management items.

Relations to other variables

For the convergent validity in study II and the supplementary analyses in this summary, it was found that the MISA-DK total scale correlates adequately and significantly to constructs related to "ingestion" (35) in terms of cognition, physical function, and orofacial function, but less to swallowing function. In study II, the multivariate regressions revealed that the variance of the MISA-DK total and subscale scores was explained more by cognitive and physical functions than of orofacial and swallowing functions. Although, it is recognized that impairments of body functions, such as orofacial and swallowing function (23), cannot predict actual occupational performance in daily life activities (37,135), our findings may raise concern of whether the items in the MISA-DK are representative for the entire construct of ingestion (35), i.e. the content validity is insufficient (56,59,60). Whether this applies to the Canadian MISA is unsolved as Lambert et al. (48) did not investigate convergent validity of these aspects. For the convergent variables in terms of cognitive function, Lambert et al. (48) found that the MISA total scale correlated less strongly than the MISA-DK. However, this might be ascribed to differences in the used measurement instrument (136) and/or the sample-dependency in the statistical methods within CTT (43,63,64).

For the Positioning and Self-feeding skills subscales of MISA-DK, all the hypotheses were confirmed in study II and in the supplementary analyses for this summary; which equals findings by Lambert et al. (48). However, adequate correlations did not continue for the removed item 9. This might underline its need for revisions. In terms of the Liquid ingestion, Solid ingestion and Texture management subscales addressing oropharyngeal skills, only one hypothesis was confirmed; namely the association of the Solid ingestion subscale to orofacial functions. The extended Rasch analysis on the Liquid ingestion subscale might shed light on our findings as discussed in the paragraph on the test content. The extended Rasch analyses and the supplementary analyses on the construct validity for the Texture management subscales might also shed light on our results. Firstly, as for the reliability analysis, the appropriateness of using statistical methods based on summarizing the texture management items for examining the construct validity might be questionable (70-72). When considering the individual texture management items, the correlation estimates reflect no associations (44) to orofacial or swallowing functions of five and ten items, respectively. In addi-

tion, four of these items (item37,38,42,43) do not discriminate significantly between the known-groups. The textures reflected in these items (pure and pudding) are assumed to enhance the efficiency and safety of swallowing (5,22,134). This might explain our findings, and support the above discussion on the need for reconsiderations and revisions of the texture management items.

The known-group comparisons of the MISA-DK subscales and total scores in study II revealed that they discriminated significantly among frail patients versus robust patients. It was also found that patients with pneumonia obtained significantly lower scores in liquid and solid ingestion versus patients without pneumonia. These findings continued during the supplementary analyses in this summary. This could reflect effects due to presbyphagia which have resulted in dysphagia (5-8,26,137). However, whether our findings reflect the presence of aspiration pneumonia or pneumonia caused by other factors (9) are unclear as we did not differentiate the aetiologies behind the pneumonia diagnoses.

Consequences of testing

The clinical utility of MISA-DK was obtained by means of expert-panel judgment and pilot testing in study I. However, information on the impact of MISA-DK in clinical practice as well as its acceptability by the patient remains to be addressed (42). In study II, no DIF by age was found, when using the sample median of 83 years. However, whether DIF would be present across more age-groups remain to be investigated. Additionally, as the MISA-DK includes similar clinical features assessing the risk of aspiration as the WST, which have been found to display high sensitivity and low specificity (138,139), an overestimation might have occurred. Finally, the greater variation in the MISA-DK scores between raters than within raters found in study III might impact on treatment planning and outcome evaluations across different therapists (36,42,43,68,118,121).

6.2. Methodological considerations

All our studies have been based on methodological research. This was deemed necessary for future research as well as for the contribution to an evidence-based occupational therapy assessment process. However, some specific methodological issues are to be addressed.

The role of the author's involvement

One general methodological limitation is related to the fact that TH has been involved in some of the data collection and all the analyses. Therefore, there is a risk of researcher bias. In order to minimize this risk, an independent research assistant performed the additional data collections.

Validity of the studies

In study I, the MISA was translated into Danish. No official guidelines could be found addressing the translation and adaptation of observation based measurement instruments, and a collaborative translation approach (57) was adapted and involved professional translators and experts within the field. This is opposite to most guidelines for translations of PROMs (140), in which non-professionals and non-experts are involved in the initial phases of the translations (58). Since the MISA is an observation based measurement instrument with explicit instructions and scoring descriptions for therapists, a translation requires good language skills and knowledge of profession-specific vocabulary, and does not require to be understood by the general population (140).

The psychometric properties of the measurement instruments used for the convergent validation in study II had been questioned (138,139,141,142), which might have influenced our results. In addition, it cannot be excluded that the convergent validity of the MISA-DK subscales and items addressing swallowing functions would have been confirmed more strongly if trial swallows using different viscosities (138) were included. In study II, the operational definitions of the frailty criteria differed from Fried et al. (4) in terms of exhaustion, which we measured using the WHO-5 and the reduced physical activity which we measured by a BI score <50. However, comparable modifications have been implemented in other studies (143).

Statistical conclusions

In study I, the CVI (88,103,105) and the AD index (104) were applied in order to quantify the experts endorsement of the content validity domains of the MISA-DK. Thirteen experts were included, which exceeds the required maximum number of ten suggested by Lynn (88). In addition, a universal agreement approach was considered with the requirement of a CVI=1 (105). However, the precession of any estimates is a function of the sample size (44,105,106), and in order to obtain a high degree of agreement with a high degree of confidence, a larger number of experts would have been beneficial.

For the Rasch analysis in study II, the Person-item-threshold distribution revealed a slightly skewed sample when analysing the MISA-DK total scale and a high percentage of extreme scores when analysing the individual subscales. This resulted in very low PSIs, although Cronbach's alpha was relatively high and constant. This reflects suboptimal targeting (102,144), which results in decreased estimates precision (79,115). Therefore, replications in larger and better targeted samples with lower levels of ingestive skills are needed. In the extended Rasch analyses, the unidimensionality t-tests of the subscales were generally adequate. However, there is an issue of power when

relatively few items/thresholds are involved in the comparisons (111). In study II, the known-groups validity was confirmed for the Liquid ingestion subscale. However, when adjusting LID during the extended Rasch analysis, the PSI, and thus the reliability, decreased to a non-sufficient level (108). Therefore, the result might be questionable.

In study III, the calculation of the $ICC_{1.3}$ used the absolute agreement definition. This coincides $ICC_{1.3}$ with the consistency definition in case of no systematic differences between the repeated measurements (65,68), and was evident for our data (unpublished observations). For the item level reliabilities, PO and Kw using quadratic weights were applied. A paradox of Kappa is its dependency on the prevalence and the marginal distributions (116,122). This was reflected in our data as some items obtained good to excellent PO, but poor to fair Kw estimates. In addition, when Kappa is calculated for non-unique pairs of raters, the 95% CI might be overestimated (123). In the interpretation of Kappa, the criteria by Cicchetti et al. (116) were used. They consider reliability in terms of clinical applications rather than research; hence, the upper levels are somewhat more stringent than other suggested criteria (106,124). However, all criteria have a level of arbitrariness (44,116,122).

Generalizability of the studies

The patient sample might not be representative for acutely hospitalised elderly medical patients since only about 25% of 439 eligible patients were included. In addition, the Person-Item threshold distributions in study II revealed that the sample did not show the low levels of ingestive skills. Furthermore, if the MISA-DK is to be administered among patients who differ from our study sample, it can be argued that new reliability testing is needed because of the sample dependency of the reliability statistics within CTT (43,63-66). In study III, the large number of raters might have influenced our results and fewer would have been preferable. This was not realizable, and in clinical practice it is not given that the same limited sets of therapists provide services to the patients. In that sense, our results may reflect the clinical reality in which the MISA-DK is to be implemented.

7. Conclusion, implications and perspectives

Prior to this PhD study, a literature review concluded that the MISA possessed adequate evidence on validity and reliability, and it was hypothesised that it could be used by occupational therapists in a Danish context. The implications of this PhD study are related to the documentation of the psychometrical properties of MISA-DK from a CTT perspective and from a Rasch model perspective,

which provided complementary information, and the following conclusions can be drawn:

- By means of expert panel judgments and linking of the MISA-DK to the ICF in study I, it was found that overall, the content of the MISA-DK clearly and adequately reflects occupational performance in eating and drinking. However, the operational definitions of the 13 items in the two texture management scales appeared ambiguous which was confirmed by a high content density ratio. In addition, the Rasch analyses in study II revealed that the item response categories of these items do not operate as intended and extended Rasch analysis revealed that the items within these two scales do not share a common underlying dimension. Using CTT in study II and supplementary analysis in this summary, the convergent validity of the Texture management items was not supported as well. This implies that these items in their current form are to be regarded as single items. In addition, reconsideration of their purpose and major revisions might be necessary.
- By means of CTT, it was found that the MISA-DK total scale, subscales and the majority of the texture management items discriminate relevantly and significantly between known-groups in terms of frailty status and pneumonia. The MISA-DK total scale converged to cognitive, physical and orofacial functions, reflecting the complexity of occupational performance in eating and drinking. The MISA-DK subscales focusing on pre-oral functions presented excellent convergent validity and equals the Canadian MISA. However, the subscales focusing on oropharyngeal functions converged only partially to measures of oropharyngeal functions. For the Liquid ingestion subscale, the extended Rasch analysis revealed that the items in this scale are not representative for the underlying construct. Although convergent variables related to oropharyngeal functions have not been addressed for the Canadian MISA, this implies a need for developing and adding more items representing oropharyngeal skills for Liquid ingestion.
- The extended evaluation of the validly of the MISA-DK using Rasch analysis in study II and an extended Rasch analyses in this summary, revealed that substantial local item dependency among items within each subscale was present. In addition, it was necessary to remove one item because of non-uniform DIF and all of the texture management items as they did not obtain fit to the model. As local item dependency might be caused by response dependency or multidimensionality, this implies further validation using factor analytic methods in conjunction with more complex Rasch models. In addition, revisions of the excluded items are needed.
- By means of CTT, relative inter- and intra-rater reliability of the original and Rasch revised MISA-DK subscale and total scale were found good to excellent in study III, which equals the

Canadian MISA. The extended evaluation of the reproducibility of the MISA-DK in terms of absolute reliability revealed that greater measurement errors were present between raters than within raters, and the supplementary reliability analysis of the individual items found good to excellent reliability estimates for 15 items between raters and for 41 items within raters. This implies that comprehensive training in the administration of MISA-DK is required in order to improve interrater reliability, and review and possible revisions of the least reliable items are needed.

7.1. Implication for clinical practise and research

This PhD study illustrates that the MISA-DK is not completely ready to be used in clinical practise or research. It seems that the conceptualization of the construct "ingestion" in relation to the texture management items has to be reconsidered, and the purpose of the subscales has to be revised. It could be suggested that the textures are regarded as different meal-task challenges. In order to verify this, validation using more complex Rasch models such as a many-facet Rasch model (79) could be suggested. Thus, parameters on both item difficulty and meal-task challenges will be included (79). Additionally, further investigation of its dimensionality (43,44,63,110,131) is needed. In the long term, assessment and adjustment of rater severity using the many-facet Rasch model (79,133) are to be included in the validation. Analyses of the sensitivity and specificity of the MISA-DK items related to swallowing safety and efficiency have to be performed using the VFS or the fiber-optic endoscopic examination of swallowing as gold standards (25,38). Also DIF analyses across more age-groups and across different diagnoses associated with dysphagia (5-8,10-13,22,30) are very relevant. If necessary, then norms are to be developed. As the MISA-DK addresses functional performance in a natural mealtime context, it might add important information in intervention studies on the efficacy of dysphagia management strategies (5-7) as well as in cohort studies on the associations of the development of frailty and dysphagia (30). As such, MISA-DK has to be invariant by different time points, which also requires DIF analyses (55,64).

This PhD study points out a dilemma in relation to Copyright agreements when translating and adapting a measurement instrument, which constraints the possibility of radical changes. However, initial Rasch analysis on data obtained with the Canadian MISA has revealed similar results as for the MISA-DK (unpublished observations). Therefore, the above suggested revisions, also apply to the Canadian MISA. After these revisions, a large cross-national study investigating whether the validity of both language versions have improved and whether they behave invariantly is needed in order to fully establish functional equivalence (50,51,114).

English abstract

Dysphagia in frail elderly patients, from an occupational therapy perspective: Danish translation and validation of the McGill Ingestive Skills Assessment for observation-based measurement of occupational performance in eating and drinking activities.

The overall purpose of this thesis was to produce a valid and reliable Danish version of the Canadian "McGill Ingestive Skills Assessment (MISA), for observational measurement of frail elderly dysphagic patients' occupational performance in eating and drinking during a meal. MISA contains 43 ingestive skills items distributed in six subscales: Positioning, Self-feeding skills, Liquid ingestion, Solid ingestion, Texture management liquids and Texture management solids. All items are scored on a 3-point ordinal scale, which are summed into subscales- and a total score. Three studies were conducted and constitute the three papers of the thesis. In addition, supplementary statistical analyses were conducted and presented in the summary of this thesis.

Methods: In order to obtain conceptual and semantic equivalence, the MISA was translated into a Danish version (MISA-DK) via a comprehensive translation procedure, inclusive judgment by experts (n=13) and pilot-test by occupational therapists (n=16). The content validity was further examined via linking of MISA-DK to the categories in the "International Classification of Function, Disability and Health" (ICF). To evaluate the validity and reliability of the MISA-DK, data were collected via two designs. The MISA-DK was administered to elderly acute medical in-patients (n= 110) as in-person observation in a prospective, consecutive, cross-sectional design. Data on external validity variables were collected in order to evaluate the convergent and known-groups validity. In addition, the patients (n=102) were video-recorded during the meal, and the video-recordings were integrated into a two-rater and test-retest design evaluating the rater reliability amongst 38 special educated raters (occupational therapists). Data were analysed using statistical methods within item response theory (i.e. the Rasch model) and classical test theory.

Results: The expert judgment and pilot-testing indicated that the content of the MISA-DK, in general, was adequate, clear, and relevant, but the items in the two texture management subscales appeared ambiguous, which was confirmed by a high content density ratio. The content of MISA-DK was related to relevant ICF categories, although the content diversity was low. The MISA-DK total

scale met the requirements of the Rasch model after adjustment of substantial local dependency between items within each subscale. Rasch analysis of the individual six subscales revealed that it was possible to achieve fit to the model for four scales; although local item dependency inflated the reliability of the Liquid ingestion subscale. The two texture management subscales did not succeed adequate fit to the Rasch model, and their items were considered as single items. Analyses within classical test theory before and after the Rasch analyses revealed, that the internal consistency reliability was adequate for the MISA-DK subscales, but relatively high for the total scale. The convergent validity was supported for the Positioning- and Self-feeding skills subscales and partially supported for the Solid ingestion subscale and the MISA-DK total scale. It was not possible to establish convergent validity in terms of orofacial and swallowing functions of the Liquid ingestion subscale and the texture management items. The known-groups validity of the MISA-DK sub- and total scales was confirmed, in that frail patients showed significantly lower ability levels within all subscales versus robust patients. Patients with pneumonia presented significantly lower ability levels in ingestion of liquid and solid foods versus patients without pneumonia. The MISA-DK demonstrated good to excellent inter- and intra-rater reliability. The amount of measurement error was small for the MISA-DK total scale, but relatively large for the subscales between raters. Reliability analyses of the 43 item using weighted Kappa statistic indicated good to excellent interrater reliability for 15 items and good to excellent intrarater reliability for 41 items.

Conclusion: When using statistical methods within classical test theory, the MISA-DK possesses adequate psychometrical properties relative to the Canadian MISA by means of convergent and known-groups validity and rater reliability. However, using the Rasch model revealed that the two texture management subscales did not met the requirements of the model and local item dependency was an evidently feature of all the MISA-DK subscales, which inflated the reliability. Thus, summation of the 43 MISA-DK items into a total score is not a valid measure of patients' ingestive skill ability during a meal. This suggests that before the MISA-DK is implemented into clinical practise and research, the texture management subscales are revised, more items reflecting additional aspects of ingestive skills ability are added and more complex Rasch models are applied for further validation and parameter estimation. Additionally, in order to improve the interrater reliability, revisions of some items and comprehensive training in the administration of the MISA-DK are recommended.

Dansk resumé

Dysfagi hos skrøbelige ældre patienter, set fra et ergoterapeutisk perspektiv: Dansk oversættelse og validering af McGill Ingestive Skills Assessment til observationsbaserede måling af aktivitets-udførelse i spise og drikke aktiviteter.

Hovedformålet med Ph.d. studiet var at udarbejde en valid og reliabel dansk version af den canadiske "McGill Ingestive Skills Assessment (MISA) til observationsbaseret måling af ældre skrøbelige dysfagi patienters aktivitetsudførelse ved indtagelse af mad og drikke under et måltid. MISA indeholder 43 items inddelt i seks underskalaer: siddestilling; spise- og drikkefærdigheder; indtagelse af væske; indtagelse af fast føde; konsistenshåndtering-væske og konsistenshåndtering-fast føde. Alle items scores på en tredelt ordinal skala, der opsummeres indenfor hver underskala og i én totalscore. Der blev gennemført tre studier, der udgør afhandlingens tre artikler. Derudover indeholder afhandlingen supplerende statistiske analyser.

Metode: Med henblik på at opnå konceptuel og semantisk ækvivalens, blev MISA oversat til dansk (MISA-DK) via en omfattende oversættelsesprocedure, inklusiv vurdering af eksperter (n = 13) og pilottest af ergoterapeuter (n = 16). Indholdsvaliditeten blev yderligere undersøgt via en sammenkædning af MISA-DK til kategorierne i "International Klassifikation af Funktionsevne, Funktionsevnenedsættelse og Helbredstilstand" (ICF). Validiteten og reliabiliteten af MISA-DK blev evalueret med data indsamlet via to designs. MISA-DK blev udført som direkte observation af ældre medicinske akut-indlagte patienter (n = 110) i et prospektivt, konsekutivt, tværsnits-design. Data på eksterne validitetsvariabler blev indsamlet med henblik på at evaluere konvergent og known-groups validitet. Desuden blev patienterne (n = 102) filmet under måltidet med video, og videooptagelserne blev integreret i et to-rater og test-retest design med henblik på at evaluere inter- og intra-tester reliabiliteten blandt 38 specialuddannede bedømmere (ergoterapeuter). Data blev analyseret med statistiske metoder indenfor item responsteori (dvs. Rasch-modellen) og klassisk testteori.

Resultater: Ekspertvurderingen og pilot-testen viste, at indholdet af MISA-DK generelt var adækvat, klart og relevant, men at items indenfor de to underskalaer for konsistenshåndtering forekom tvetydige, hvilket blev bekræftet ved en høj indholdsdensitetsratio. Indholdet af MISA-DK var relateret til relevante ICF kategorier, dog var indholdsdiversiteten lav. Den samlede MISA-DK skala opfyldte kravene i Rasch-modellen efter justering af betydelig lokal afhængighed mellem items

indenfor hver underskala. Raschanalyser af de individuelle underskalaer viste at fire opfyldte kravene i modellen; dog betød lokal item afhængighed i underskalaen for indtagelse af væske, at reliabiliteten var kunstig høj for denne skala. Items indenfor de to underskalaer for konsistenshåndtering opfyldte ikke kravene i Rasch-modellen, og deres score på den tredelte ordinal skala bør ikke opsummeres. Analyser indenfor klassisk testteori før og efter Rasch analyserne, viste at intern konsistens reliabiliteten var acceptabel for de seks underskalaer, men relativ høj for den samlede skala. Konvergent validitet blev bekræftet for to underskalaer (siddestilling samt spise- og drikkefærdigheder), og blev delvist bekræftet for én underskala (indtagelse af fast føde) og den samlede MISA-DK skala. Det var ikke muligt at bekræfte konvergent validitet for tre underskalaer (indtagelse af væske, konsistenshåndtering-fast føde og konsistenshåndtering-væske). Known-groups validitet blev bekræftet, idet skrøbelige patienter præsenterede et signifikant lavere færdighedsniveau indenfor alle underskalaerne i sammenligning med robuste patienter. Patienter med lungebetændelse præsenterede et signifikant lavere færdighedsniveau ved indtagelse af væske og af fast føde i sammenligning med patienter uden lungebetændelse. MISA-DK demonstrerede god til fremragende inter- og intra-tester reliabilitet. Standardmålefejlen var generelt lav for den samlede MISA-DK score men var relativ høj for underskalaerne mellem bedømmere. Reliabilitetsanalyse af de individuelle 43 items med vægtet Kappa statistik viste at 15 items demonstrerede god til fremragende inter-tester reliabilitet og 41 items demonstrerede god til fremragende intra-tester reliabilitet.

Konklusion: Når statistiske metoder indenfor klassisk testteori blev benyttet, besidder MISA-DK adækvate psykometriske egenskaber relativt til den canadiske MISA med hensyn til konvergent og known-groups validitet og tester-reliabilitet. Dog viste analyser med Rasch modellen, at de to underskalaer for konsistenshåndtering ikke opfylder kravene i modellen og lokal item afhængighed var et evident træk for alle MISA-DK underskalaerne, hvilket betød at reliabiliteten var kunstig høj. Derfor er en opsummering af de 43 MISA-DK items til én samlet score ikke et validt mål for patienters aktivitetsudførelse ved indtagelse af mad og drikke under et måltid. Det betyder at: underskalaerne for konsistenshåndtering bør revideres; flere items, der afspejler supplerende aspekter af aktivitetsudførelse ved indtagelse af mad og drikke bør tilføjes; og mere komplekse Rasch modeller bør anvendes til yderligere validering og parameter estimering inden MISA-DK implementeres i klinisk praksis og forskning. For at forbedre inter-tester reliabiliteten, anbefales det at enkelte items revideres samt at den enkelte terapeut uddannes grundigt i brugen af MISA-DK.

Acknowledgments

The completion of this thesis has only been possible with the support of many people. I would especially like to thank:

Jens Faber, my main supervisor, for giving engaged, encouraging and inspiring support from the very beginning of the study and to the completion of this thesis. Thank you for your incredibly open-mindedness and desire to gain insight into the occupational therapy profession and way of reasoning, and to help me focusing on the hypothesis and the research problems, see through weaknesses, and to maintain the "red thread" in my research. Thank you for your belief in me.

Heather Lambert, my co-supervisor, for her huge experience in the field of dysphagia, occupational therapy and psychometrics, for never failing to provide support and supervision either by mail correspondence or in-person when I visited the McGill University in Montréal, and for letting me develop a Danish version of "the McGill Ingestive Skills Assessment".

Trine Pedersen, my co-supervisor, for continuously engaged support through the data collection in a busy working day at the ward.

Charlotte Ehlers Hansen, my research assistant, for careful data collection for study II.

All participating patients, for letting us use their time and energy and for allowing us to observe and video-record their performance during a meal, which to a great extend is a very private matter.

Participating occupational therapists: Camilla, Irene, Tanja, Anders, Evelyn, Lea, Dorthe, Lotte, Anne, Mine, Marianne, Rikke, Anita, Mette, Hanne, Jette, Lilly, Nathalie, Kaya, Jesper, Pernille, Mette, Claus, Signe, Mia, Tine, Charlotte, Tina, Jette, Lise, Heidi, Kari, Lasse, Julie, Birgitte, Annette, Trine, Elisabeth, Rikke, Stine, Elsa-Maria, Marianne, Brit and Sanne, for contribution to study I and III.

Ton Satink my former supervisor at the Academy for European Master's Degree Study in Occupational Therapy, for encouraging me to continue my academic education and to do research.

Svend Kreiner at the university in Copenhagen, for opening the door to the universe of Georg Rasch and for giving me faith that psychometric testing is as scientific as testing of the effects and side effects of medicine.

Alan Tennant and Mike Horton from the Psychometric Laboratory for Heath Science, University of Leeds, for organizing a highly pedagogical work-shop in the application of the Rasch meas-

urement model.

Inja Thoustrup, Hanne F. Skall and all staff members in the occupational therapy department, for providing opportunities and conditions for my scientific work, and for continuous support.

Merete Wormslev, my office mate and colleague, for inspiring discussions on the ICF as well as on didactics, and for giving careful support when needed.

Annette Kjærsgaard and staff members at Hammel Neurocenter, for giving me the opportunity to learn from their huge knowledge and experience within the field of dysphagia and for fruitful discussions on evidence-based measurement instruments of dysphagia within occupational therapy.

My husband Peter, for giving tolerant, engaged, loving and caring support throughout the work. Without you, it would not have succeeded. My daughter Maja, for giving very pedagogical "Special Education" in mathematics. My family and friends, for continued support all the way.

Financial support for this thesis has been received from Herlev Hospital Research Council, the Occupational Therapy Association Research Foundation and the Lundbeck Foundation. Thanks for your support

Tina Hansen
Solrød Strand, 2012

References

1. Befolkningsfremskrivninger 2011-2050. I: Nyt fra Danmarks Statistik nr. 205 maj 2011. København, Danmarks Statistik. Available from: http://www.dst.dk/pukora/epub/Nyt/2011/NR205.pdf.
2. Sygehusbenyttelse 2010. I: Nyt fra Danmarks Statistik nr. 363 december 2011. København, Danmarks Statistik. Available from: http://www.dst.dk/pukora/epub/Nyt/2011/NR636.pdf.
3. Bergman H, Ferrucci L, Guralnik J, Hogan DB, Hummel S, Karunananthan S, et al. Frailty: an emerging research and clinical paradigm - issues and controversies. J Gerontol A Biol Sci Med Sci 2007;62(7):731-737.
4. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56:M146-M1567.
5. Rofes L, Arreola V, Almirall J, Cabré M, Campins L, García-Peris P, et al. Diagnosis and management of oropharyngeal dysphagia and its nutritional and respiratory complications in the elderly. Gastroenterol Res Pract 2011; Article ID 818979, (Epub 2010 Aug 3).
6. Miller N, Carding P. Dysphagia: implications for older people. Rev Clin Gerontol 2007;17:177-190.
7. Schindler JS, Kelly JH. Swallowing disorders in the elderly: State of the art review. Laryngoscope 2002;122:589–602.
8. Achem SR, DeVeault KR. Dysphagia in aging. J Clin Gastroentero 2005;39:357-371.
9. van der Maarel-Wierink CD, Vanobbergen JNO, Bronkhorst EM, Schols JMGA, de Baat C. Risk factors for aspiration pneumonia in frail older people: a systematic literature review. J Am Med Dir Assoc 2010;12(5):344-354.
10. Altman KW, Yu GP, Schaefer SD. Consequences of dysphagia in the hospitalized patient: Impact on prognosis and hospital resources. Arch Otolaryngol Head Neck Surg 2010;136(8):784-789.
11. Cabré M, Serra-Prat M, Palomera E, Almirall J, Pallares R, Clavé P. Prevalence and prognostic implications of dysphagia in elderly patients with pneumonia. Age Ageing 2010;39(1):39-45.
12. Gayomard V, Fulcher RA, Redmayne O, Metcalf AK, Potter JF, Myint PK. Effect of dysphasia and dysphagia on inpatient mortality and hospital length of stay: A database study. J Am Geriatr Soc 2009;57:2101-2106.
13. Lee A, Sitoh YY, Lieu PK, Phua SY, Chin JJ. Swallowing impairment and feeding dependency in the hospitalized elderly. Am Acad Med Singapore 1999;28:371-376..

14. Clark GF, Avery-Smith W, Wold LS, Anthony P, Holm SE. Eating and Feeding Task Force; Commission on Practice (AOTA). Specialized knowledge and skills in feeding, eating, and swallowing for occupational therapy practice. Am J Occup Ther 2007;61(6):686-700.
15. Medin J, Larson J, Von Arbin M, Wredling R, Tham, K. Elderly persons' experience and management of eating situations 6 month after stroke. Disabil Rehabil 2010;32:1346-1353.
16. Perry L, McLaren S. Coping and adaptation at six months after stroke: experiences with eating difficulties. Int J Nurs Stud 2003;40:185-195.
17. Perry L, McLaren S. An exploration of nutrition and eating disabilities in relation to quality of life at 6 months post-stroke. Health Soc Care Comm 2004;12:288-297.
18. McHorney CA, Robbins J, Lomax K, Rosenbek JC, Chignell K, Kramer AE, et al. The SWAL-QOL and SWAL-CARE outcomes tool for oropharyngeal dysphagia in adults, III: Documentation of reliability and validity. Dysphagia 2002;17:97–114.
19. Chen PH, Golub JS, Hapner ER, Johns MM. Prevalence of perceived dysphagia and quality-of-life impairment in a geriatric population. Dysphagia 2009;24(1):1-6.
20. Miura H, Yamasaki K, Morizaki N, Moriya S, Sumi Y. Factors influencing oral health-related quality of life (OHRQoL) among the frail elderly residing in the community with their family. Arch Gerontol Geriatr 2010;51(3):e62-e65.
21. Eslick GD, Talley NJ. Dysphagia: epidemiology, risk factors and impact on quality of life- a population-based study. Aliment Pharmacol Ther 2008;27(10):971-979.
22. Cichero JYA, Murdoch BE. Dysphagia: Foundation, theory and practice. Chichester: John Wiley & Sons Ltd; 2006.
23. Threats TT. Use of the ICF in dysphagia management. Semin Speech Lang 2007; 28:323-333.
24. Lambert HC, Gisel EG, Wood-Dauphine S, Groher ME, Abrahamowicz M. McGill Ingestive Skills Assessment: User's manual. Ottawa: Canadian Association of Occupational Therapists; 2006.
25. Hansen T, Kjærsgaard A, Faber J. Measuring elderly dysphagic patients' performance in eating - a review. Disabil Rehabil 2011;33(21-22):1931-1940.
26. Humbert IA, Robbins J. Dysphagia in the elderly. Phys Med Rehabil Clin N Am 2008;(19):853-866.
27. Roy N, Stemple J, Merrill RM, Thomas L. Dysphagia in the elderly: preliminary evidence of prevalence, risk factors, and socioemotional effects. Ann Otol Rhinol Laryngol 2007;116(11):858-865.

28. Serra-Prat M, Hinojosa G, López D, Juan M, Fabré E, Voss DS, et al. Prevalence of oropharyngeal dysphagia and impaired safety and efficacy of swallow in independently living older persons. J Am Geriatr Soc 2011;59(1):186-187.
29. Holland G, Jayasekeran V, Pendleton N, Horan M, Jones M, Hamdy S. Prevalence and symptom profiling of oropharyngeal dysphagia in a community dwelling of an elderly population: a self-reporting questionnaire Survey. Dis Esophagus 2011;24:476-480.
30. Rofes L, Arreola V, Romea M, Palomera E, Almirall J, Cabré M, et al. Pathophysiology of oropharyngeal dysphagia in the frail elderly. Neurogastroenterol Motil 2010;22:851–859.
31. Miura H, Kariyasu M, Yamasaki K, Arai Y. Evaluation of chewing and swallowing disorders among frail community-dwelling elderly individuals. J Oral Rehabil 2007;34:422-427.
32. Lin LC, Wu SC, Chen HS, Wang TG, Chen MY. Prevalence of impaired swallowing in institutionalized older people in Taiwan. J Am Geriatr Soc 2002;50(6):1118-1123.
33. World Health Organization. International Classification of Functioning, Disability and Health (ICF). Geneva: World Health Organization; 2001.
34. Swigert NB. Update on current assessment practices for dysphagia. Top Geriatr Rehabil 2007;23:185–196.
35. Leopold NA, Kagel MC. Dysphagia--ingestion or deglutition? A proposed paradigm. Dysphagia 1997;12(4):202-206.
36. Dunn W. Measurement issues and practices. In: Law M, Baum C, Dunn W, (Eds). Measuring occupational performance: Supporting best practice in occupational therapy. 2nd ed. Thorofare, NJ: Slack; 2005.
37. Fisher AG. Occupational Therapy Intervention Process Model: A model for planning and implementing top-down, client-centred, and occupation-based interventions. Fort Collins, CO: Three Star Press; 2006.
38. Meriano C, Latella D. Occupational therapy interventions: Function and occupation. Thorofare, NJ: Slack; 2008.
39. Kjærsgaard A. Undersøgelse af synkeproblemer hos senhjerneskadede patienter. (Assessment of swallowing difficulties in brain injured patients). SATS-puljeprojektet. 2002. Available from URL:http://www.vfhj.dk/admin/write/files/824.pdf.
40. Canadian Association of Occupational therapy. CAOT Position Statement: Feeding, eating and swallowing and occupational therapy. 2010. Available from: URL: http://www.caot.ca/default.asp?pageid=3948.

41. Kjærsgaard A. Ansigt, mund og svælg: Undersøgelse og behandling efter Coombes konceptet. (Face, mouth and throat: assessment and treatment a.m. the Coombes concept). Copenhagen: FADL; 2005.
42. Laver Fawcett A. Principles of assessment and outcome measurement for occupational therapists and physiotherapists: theory, skills and application. Chichester: John Wiley & Sons Ltd; 2007.
43. Streiner D, Norman G. Health measurement scales. 3rd ed. Oxford: Oxford University Press; 2003.
44. Portney IG, Watkins MP. Foundations of clinical research. Application to practice. 3rd ed. Upper Saddle River, NJ: Pearson & Prentice Hall; 2009.
45. Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable? Injury 2011;42:236-240.
46. Dekker J, Dallmaijer AJ, Lankhorst GJ. Clinimetrics in rehabilitation medicine: Current issues in developing and applying measurement instruments 1. J Rehabil Med 2005;37:193–201.
47. Lambert HC, Gisel EG, Groher ME, Wood-Dauphine S. McGill Ingestive Skills Assessment (MISA) Development and first field test of an evaluation of functional ingestive skills of elderly persons. Dysphagia 2003;18:101-113.
48. Lambert HC, Gisel EG, Groher ME, Abrahamowicz M, Wood-Dauphine S. Psychometric testing of the McGill Ingestive Skills Assessment. Am J Occup Ther 2006;60:409-419.
49. Lambert HC, Abrahamowicz M, Groher ME, Wood-Dauphine SM, Gisel EG. The McGill Ingestive Skills Assessment predicts time to death in an elderly population with neurogenic dysphagia: Preliminary evidence. Dysphagia 2005;20:123-132.
50. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQOL instruments: the universalist approach. Qual Life Res 1998;7:323-235.
51. Geisinger KF. Cross-cultural normative assessment: Translation and adaptation issues influencing normative interpretation of assessment instruments. Psychol Assess 1994;4:304–312.
52. Beckstead JW, Yang CY, Lengacher CA. Assessing cross-cultural validity of scales: A methodological review and illustrative example. Int J Nurs Stud 2008;45:110–119.
53. van Raaij WF. Cross-cultural research methodology as a case of construct validity. Adv Consum Res 1978;5:693-701.
54. Brown T. Constructs validity: A unitary concept for occupational therapy assessment and measurement HKJOT 2010;20(1):30-42.

55. Smith EV. Evidence of the reliability of measures and validity of measure interpretation: A Rasch measurement perspective. J Appl Meas 2001;2(3):281-311.
56. Wilson M. Constructing measures: an item response modeling approach. Mahwah, NJ: Lawrence Erlbaum Associates; 2004 (reprinted 2010).
57. Douglas SP, Craig CS. Collaborative and iterative translation: An alternative approach to back translation. J Int Mark 2007;15:30–43.
58. McKenna S. Measuring patient-reported outcomes: moving beyond misplaced common sense to hard science. BMC Medicine 2011; 9: 86.
59. Sireci SG. The construct of content validity. Social Indicators Res 1998;45:83–117.
60. Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: A functional approach to concepts and methods. Psychol Assess 1995;3:238–247.
61. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Üstun B, Stucki G. ICF linking rules: An update based on lessons learned. J Rehabil Med 2005;37:212–218.
62. Xiong T, Hartley S. Challenges in linking health-status outcome measures and clinical assessment tools to the ICF. Arch Physio 2008;10:152–156.
63. Crocker LM, Algina J. Introduction to classical and modern test theory. Mason, OH: Cengage Learning; 2006.
64. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. Health Tecnol Assess 2009:13(12).
65. Weir JP. Quantifying Test-Retest reliability using the Intraclass Correlation Coefficient and the SEM. J Strength Cond Res 2005;19:231-240.
66. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson, A, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. J Clin Epidemiol 2011;64:96-106.
67. Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. Am J Phys Med Rehabil 2005;84:719-723.
68. de Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol 2006;59:1033-1039.
69. Bruton A, Conway JH, Holgate ST. Reliability: What is it and how is it measured. Physiotherapy 2000;86:94–99.
70. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989;70:857-860.

71. Tesio L. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. J Rehabil Med 2003; 35:105–115.
72. Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. Lancet Neurol 2007;6:1094-1105.
73. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Care Res 2007;57:1358–1362.
74. Salzberger T. Does the Rasch Model Convert an Ordinal Scale into an Interval Scale? Rasch Meas Trans 2010;24:1273-75
75. Ewing MT, Salzberger T, Sinkovics RR. Confirmatory factor analysis vs. Rasch approaches: Differences and Measurement Implications, Rasch Meas Trans 2009;23: 1194-5.
76. Linacre JM. Rasch Analysis First or Factor Analysis First? Rasch Meas Trans 1998;11: 603.
77. Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago: University of Chicago Press; 1960 (reprinted 1980).
78. Kreiner S. Validity and objectivity: Reflections on the role and nature of Rasch models. Nordic Psychology 2007;59(3):268-298.
79. Bond TG, Fox CM. Applying the Rasch Model: Fundamental measurement in the human sciences. 2en ed. Mahwah, NJ: Lawrence Erlbaum; 2007.
80. Andric D. A rating formulation for ordered response categories. Psychometrika 1978;43:561–573.
81. Masters GN. A Rasch model for partial credit scoring. Psychometrika 1982;47:149–174.
82. Linacre JM. Optimizing rating scale category effectiveness. J Appl Meas 2002;3(1):85–106.
83. Andrich D, de Jong J, Sheridan B. Diagnostic opportunities with the Rasch model for ordered response categories. In: Rost J, Langeheine R (Eds). Applications of latent trait and latent class models in the social sciences. New York: Waxmann Münster; 1997.
84. Rosenbuam PR. Criterion-related construct validity. Psychometrika 1989;54(4):625-633
85. Danish ICF network via Hanne Melchiorsen, Marselisborg- Centret (Homepage on the internet) (cited 21 September 2009). Available at:ttp://www.marselisborgcentret.dk/kontakt-os/forskning-udvikling/hanne-melchiorsen/#c1569
86. Memorandum of agreement made this 24th day of September 2010 between Canadian Association of Occupational Therapists, Ottawa; Ontario, Canada and Tina Hansen, Herlev University

Hospital, Denmark.
87. Grant JS, Davis LL. Focus on quantitative methods: Selection and use of content experts for instrument development. Res Nurs Health 1997;20:269–274.
88. Lynn MR. Determination and quantification of content validity. Nurs Res 1986;35:382–385
89. Stamm T, Geyh S, Cieza A, Machold K, Kollerits B, Kloppenburg M, et al. Measuring function in patients with hand osteoarthritis-content comparison of questionnaires based on the International Classification of Function, Disability and health (ICF). Rheumatology 2006;45:1534–1541.
90. 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.
91. Kørner EA, Lauritzen L, Nilsson FM, Wang A, Christensen P, Lolk A. Mini mental state examination. Validation of new Danish version. Ugeskr Laeg 2008;170:745–749.
92. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index in stroke rehabilitation. J Clin Epidemiol 1989;42(8):703-709.
93. Bakke M, Bergendal B, McAllister A, Sjögreen L, Åsten P. Development and evaluation of a comprehensive screening for orofacial dysfunction. Swed Dent J 2007;31:75–84.
94. Smithard DG, O'Neill PA, Park C, England R, Renwick DS, Wyatt R, et al. Can bedside assessment reliably exclude aspiration following acute stroke? Age Ageing 1998;27:99-106.
95. Kondrup J, Rasmussen HH, Hamberg O, Stanga Z. Nutritional risk screening (NRS 2002): a new method based on an analysis of controlled clinical trials. Clin Nutr 2003;22:321–336.
96. Heun R, Burkart M, Maier W, Bech P. Internal and external validity of the WHO Well-Being Scale in the elderly general population. Acta Psychiatr Scand 1999;99:171-178.
97. Bohannon RW, Peolsson A, Massy-Westropp N, Desrosiers J, Bear-Lehman J. Reference values for adult grip strength measured with a Jamar dynamometer: a descriptive meta-analysis. Physiotherapy 2006;92:11–15.
98. Podsiadlo D, Richardson S. The Timed ''Up & Go'': A test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39(2):142–148.
99. Woehr DJ, Huffcutt AI. Rater training for performance appraisal: A quantitative review. J Occup Organ Psychol 1994;67:189-205.
100. Hounsgaard L, Eriksen JJ. Læring i sundhedsvæsenet. (Learning within Health Care). Copenhagen: Gyldendal; 2003.
101. Gingerich A, Regehr G, Eva KW. Rater-Based assessments as social judgments: Rethinking

the etiology of rater errors. Acad Med 2011;86(10 Suppl):S1-S7.
102. Andrich D, Lyne A, Sheridan B, Luo G. RUMM2030. Perth: RUMM Laboratory Pty; 2010.
103. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health 2007;30:459–467.
104. Burke MJ, Dunlap WP. Estimating interrater agreement with the Average Deviation Index: A user's guide. Organ Res Methods 2002;5:159–172.
105. Beckstead JW. Content validity is naught. Int J Nurs Stud 2009;46(9):1274–1283.
106. Altman DG. Practical statistics for medical research. New York, Chapman and Hall/CRC Press; 1991.
107. Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). Br J Clin Psychol 2007;46:1-18.
108. Fisher WP. Reliability Statistics. Rasch Meas Trans 1992;6:238.
109. Smith EV: Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. J Appl Meas 2002;3(2):205-231.
110. Tennant A, Pallant JF. Unidimensionality matters! (a tale of two Smiths?). Rasch Meas Trans 2006;20:1048–1051.
111. Lundgren-Nilsson Å, Tennant A. Past and present issues in Rasch analysis: The Functional Independence Measure (FIMTM) revisited. J Rehabil Med 2011;43:884-891.
112. Marais I, Andrich D. Formalizing dimension and response violations of local independence in the unidimensional Rasch model. J Appl Meas 2008;9:200-215.
113. Lee G, Brennan RL, Frisbie DA. Incorporating the testlet concept in test score analyses. Educ Meas I P 2000;19:9–15.
114. Tennant A, Penta M, Tesio L, Grimby G, Thonnard JL, Slade A, et al. 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(Suppl 1):I37-I48.
115. Linacre JM. Sample size and item calibration stability. Rasch Meas Trans 1994;7:328.
116. Cicchetti DV, Bronen R, Spencer S, Haut S, Berg A, Oliver P, et al. Rating scales of measurement, issues of reliability: Resolving some critical issues for clinicians and researchers. J Nerv Ment Dis 2006;194:557-564.
117. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies, Stat Med 1998;17:101-10.

118. Van Baalen B, Odding E, van Woensel MPC, van Kessel MA, Roebroeck ME, Stam HJ. Reliability and sensitivity to change of measurement instrument used in a traumatic brain injury population. Clin Rehabil 2006;20:686-700.
119. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–310.
120. Smith MW, MA J, Stafford RS. Bar charts enhance Bland-Altman plots when value ranges are limited. J Clin Epidemiol 2010;63:180-184.
121. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217-238.
122. Sim J, Wright CC. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. Phys Ther 2005;85:257–268.
123. Fleiss JL, Levin B, Paik MC. The Measurement of Interrater Agreement, in Statistical Methods for Rates and Proportions. 3rd ed. Hoboken, NJ: USA, John Wiley & Sons Inc.; 2004.
124. Cicchetti DV. The precession of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. J Clin Exp Neuropsychol 2001; 23:695-700.
125. Hedin PJ, McKenna SP, Meads DM. The Rheumatoid Arthritis Quality of life (RAQol) for Sweden: Adaptation and validation. Scand J Rheumatol 2008;35:117–123.
126. Hagell P, Hedin PJ, Meads DM, Nyberg L, McKenna SP. Effects of method of translation of patient-reported health outcome questionnaires: A randomized study of the translation of the Rheumatoid Arthritis Quality of life (RAQol) for Sweden. Value Health 2010;13:424-430.
127. Ndosi M, Tennant A, Bergsten U, Kukkurainen ML, Machado P, de la Torre-Aboki J, Vlieland TP, Zangi HA, Hill J. Cross-cultural validation of the Educational Needs Assessment Tool in RA in 7 European countries. BMC Musculoskelet Disord. 2011;12:110
128. Lundgren-Nilsson Å, Jonsdottir IH, Pallant J, Ahlborg G Jr. Internal construct validity of the Shirom-Melamed Burnout Questionnaire (SMBQ).BMC Public Health. 2012;12:1.
129. Latimer S, Covic T, Tennant A. Co-calibration of deliberate self-harm (DSH) behaviours: Towards a common measurement metric. Psychiatry research 2012, http://dx.doi.org/10.1016/j.psychres.2012.05.019
130. Amin L, Rosenbaum P, Barr R, Sung L, Klaassen RJ, Dix DB, Klassen A. Rasch analysis of the PedsQL: an increased understanding of the properties of a rating scale. J Clin Epidemiol 2012;65(10):1117-23.

131. Ip EH. Empirically indistinguishable multidimensional IRT and locally dependent unidimensional item response models. Br J Math Stat Psychol 2010;63: 395–416

132. Yen WM: Scaling performance assessments: Strategies for managing local item dependence. Journal of Educational Measurement 1993, 30:187-213.

133. Myford CM, Wolfe EW. Detecting and measuring rater effects using Many-Facet Rasch Measurement: Part 1. J Appl Meas 2003;4(4):386-422.

134. Atherton M, Bellis-Smith N, Cichero JAY, Suter M. Texture-modified foods and thickened fluids as used for individuals with dysphagia: Australian standardized labels and definitions. Nutr Diet 2007;64 (Suppl.):53 – 76.

135. Vandervelde L, Van den Bergh PYK, Renders A, Goemans N, Thonnard JL. Relationships between motor impairments and activity limitations in patients with neuromuscular disorders. J Neurol Neurosurg Psychiatr 2009;80:326–332.

136. Tombaugh T, McDowell I, Kristjansson B, Hubley AM. Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): a psychometric comparison and normative data. Psychol Assess 1996;8:48–59.

137. Hansen T, Lambert HC, Faber J. Ingestive skill difficulties are frequent among acutely-hospitalized frail elderly patients, and predict hospital outcomes. Phys Occup Ther Geriatr. Early Online: 1–17, 2012. DOI: 10.3109/02703181.2012.736019. Available online at http://informahealthcare.com/potg

138. Bours GJ, Speyer R, Lemmens J, Limburg M, de Wit R. Bedside screening tests vs. videofluoroscopy or fiberoptic endoscopic evaluation of swallowing to detect dysphagia in patients with neurological disorders: systematic review. J Adv Nurs 2009;65(3):477-93.

139. Matino R, Pron G, Diamant N. Screening for oropharyngeal dysphagia in stroke: Insufficient evidence for guidelines. Dysphagia 2000;15:19–30.

140. Schuster C, Hahn S, Ettlin T. Objectively-assessed outcome measures: a translation and cross-cultural adaptation procedure applied to the Chedoke McMaster Arm and Hand Activity Inventory (CAHAI). BMC Med Res Methodol. 2010; 10:106

141. de Morton NA, Keating JL, Davidson M. Rasch analysis of the Barthel index in the assessment of hospitalized older patients after admission for an acute medical condition. Arch Phys Med Rehabil 2008;89:641–647.

142. Schultz-Larsen K, Kreiner S, Lomholt RK. Mini-Mental Status Examination: mixed Rasch model item analysis derived two different cognitive dimensions of the MMSE. J Clin Epidemiol

2007;60:268–279.
143. Romero-Ortuno R, Walsh CD, Lawlor BA, Kenny RA. A frailty instrument for primary care: findings from the Survey of Health, Ageing and Retirement in Europe (SHARE). BMC Geriatr 2010;10:57.
144. Andrich D. An index of person separation in latent trait theory, the traditional KR20 index, and the Guttman scale response pattern. Educ Res Perspect 1982;9(1): 95-104.

Appendix A - Translation report

MISA – The McGill Ingestive Skills Assessment (Original English version). Lambert HC, Gisel EG, Wood-Dauphine S, Groher ME, Abrahamowicz M. McGill Ingestive Skills Assessment: User's manual and Evaluation forms. Copyright © 2006 by Canadian Association of Occupational Therapists. Published under arrangement with Canadian Association of Occupational Therapists. Ottawa, Ontario, Canada.
MISA-DK (Dansk projektversion). Danish language translation Copyright © 2010 by Tina Hansen, Ergoterapiafsnittet 53P1, Herlev Universitets Hospital, Region H, Herlev Ringvej 75, 2730 Herlev, Danmark.

MISA, der benyttes til observationsbaseret måling af dysfagipatienters færdigheder ved indtagelse af mad og drikke under et måltid, består af en 42 siders manual samt et fire siders registreringsark med 43 items. I manualen beskrives udviklingen og testningen af MISA, anvendelsen af MISA samt de 43 items og deres specifikke scoring.

Oversættelsesfasen indbefattede:

- Indledende oversættelse via tre oversættere (uge 50, 2008 - uge 22, 2009)
- En konsensusoversættelse via en review gruppe (tre ergoterapeuter og en diætist) samt kontrol af semantisk ækvivalens via en fjerde oversætter (uge 24-36, 2009)
- Endelig projektversion via ekspertpanel vurdering (uge 38-42, 2009)

Af hensyn til pladsmæssige ressourcer i afhandlingen, rapporters oversættelsesfasen udelukkende i relation til de enkelte item termer. Item- og scorebeskrivelser gengives ikke, men refereres til hvor det er relevant. Item- og scorebeskrivelserne findes på engelsk i originalversionen, der kan købes via Canadian Association of Occupational Therapists (http://www.caot.ca), og på dansk i den endelige projektversion af MISA-DK i Appendiks B.

De vigtigste elementer fra oversættelsesfasen fremgår af følgende:

Oversættelsesrapport for: MISA- item termer samt vigtigste elementer i item- og score beskrivelser		
Original version	Første oversættelser: 3 oversættere (A, B og C) Konsensusoversættelse: Review gruppe og semantisk kontrol (D)	Endelig oversættelse efter ekspertpanel vurdering (E)
Positioning scale	A. Positioneringsskala B. Skala for positionering C. Skala for siddestilling D. Skala for siddestilling. (”Positionering” kan have flere betydninger- ”siddestilling” vælges).	E. Skala for siddestilling. *Kommentar: Det er uklart af hvem og hvornår korrektioner af patientens siddestilling må foretages.* I skalabeskrivelsen er tilføjet: ”Hvis ergoterapeuten er eneste sundhedsprofessionelle til stede og patienten ikke kan opretholde en hensigtsmæssig siddestilling i forhold til at spise og drikke, så kan ergoterapeuten intervenerer. Dette skal dog afspejles i de givne scorer”.
1. Maintain symmetry of posture	A. Fastholder symmetri i kropsholdningen B. Opretholder symmetri i kropsholdningen C. Opretholder symmetrisk siddestilling D. Opretholder symmetrisk kropsstilling. (”Kropsstilling” benyttes i den danske ICF kap.4 aktivitet/deltagelse og vælges)	E. Opretholder symmetrisk kropsstilling. CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
2. Maintain adequate head position for feeding	A. Fastholder en passende hovedposition for spisning/ved at spise B. Opretholder passende hovedstilling for spisning C. Opretholder adækvat hovedstilling under indtagelse af mad og drikke D. Opretholder passende hovedstilling i forhold til at spise og drikke. (”Feeding” har været vanskelig at oversætte. Feeding er også synonym for ”self-feeding”. Hovedets stilling har betydning for effektiviteten og sikkerheden mht. at føre mad og drikke til munden samt at synke. Definitionen af ”spise” og ”drikke” i den danske ICF kap 5 aktivitet/deltagelse integrerer begge aspekter og vælges)	E. Opretholder passende hovedstilling i forhold til at spise og drikke CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
3. Maintain 90-degree hip angle	A. Fastholder en 90-graders vinkel i hoften B. Opretholder en hoftevinkel på 90 grader C. Opretholder 90 grader vinkel i hoften D. Opretholder 90 graders hoftefleksion. (”Hoftefleksion” bruges indenfor dansk ergoterapi terminologi og vælges).	E. Opretholder 90 graders hoftefleksion CVI=1/AD<0.65 for 3 indholdsvaliditetsdomæner. CVI=0.92 for ”klar scorebeskrivelse”. *Kommentar: Hvorfor hoftefleksion -Det er vel patientens evne til at opretholde alignment i truncus, så hovedstilling ikke er i ekstension.*

		I scorebeskrivelsen er tilføjet: (a)"I den siddende stilling ved indtagelse af mad og drikke, bør bækkenet være fremadkippet således at hoften er flekteret svt. 90 grader", og (b) Noter altid hvilket lejringshjælpemiddel patienten bruger, også selvom han opretholder 90 graders hoftefleksion".
4. Able to sit upright without leaning on arm	A. At kunne sidde oprejst uden at støtte på armen B. Opretholder siddestilling uden at læne sig på armen(e) C. Er i stand til at sidde i opret stilling uden at støtte sig på armen D. Kan sidde opret uden at støtte sig med armen	E. Opretholder postural stabilitet i truncus CVI=1/AD<0.65 for 3 indholdsvaliditetsdomæner. CVI=0.92 for "adækvat item term". *Kommentar: Item term afspejler ikke hvad item undersøger*. Item term ændret mhp. at reflektere item indhold.
Self-feeding skills scale	A. Evnen til at spise skala B. Skala for evnen til at spise selv C. Skala for færdigheder i forbindelse med indtagels af mad og drikke D. Skala for spise- og drikkefærdigheder ("Self-feeding" oversættes som i item 2).	E. Skala for spise- og drikkefærdigheder
5. Able to grasp utensil functionally and bring it to the mouth	A. At kunne gribe funktionelt fat om redskabet og bringe det op til munden B. Griber funktionelt om spiseredskaber og fører dem til munden C. Er i stand til at gribe funktionelt om bestik og fører det til munden D. Kan tage funktionelt fat om bestik/fødeemne og føre det til munden. (Da ikke al mad indtages med bestik, er "fødeemne" tilføjet. I itembeskrivelse tilføjes, at det dog bør sikres at måltidet også indeholder mad, der skal spises med bestik).	E. Kan tage funktionelt fat om bestik/fødeemne og føre det til munden CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner. *Kommentar: Uklart om hvilken grad af hjælp, der må gives i score 2 sammenlignet med score 1*. Der tilføjes, at der scores 1, hvis patienten fysisk guides og scores 2 hvis patienten instrueres.
6. Able to grasp cup/glass functionally and bring it to the mouth	A. At kunne gribe funktionelt fat om en kop/glas og føre det til munden B. Griber funktionelt om kop/glas og fører dem til munden C. Er i stand til at gribe funktionelt om kop/glas og føre det til munden D. Kan tage funktionelt fat om kop/glas og føre det til munden	E. Kan tage funktionelt fat om kop/glas og føre det til munden CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner. *Kommentar: Uklart om hvilken grad af hjælp, der må gives i score 2 sammenlignet med score 1*. Der tilføjes, at der scores 1, hvis patienten fysisk guides og scores 2 hvis patienten drikker selv, men har behov for at

		kop/glas placeres i hånden.
7. Selects appropriate utensil for food item	A. At kunne vælger det rette redskab til den pågældende mad B. Vælger passende spiseredskaber i forhold til fødeemner C. Udvælger bestik egnet til madvaren D. Vælger hensigtsmæssigt bestik i forhold til fødeemnerne	E. Vælger hensigtsmæssigt bestik i forhold til fødeemne CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
8. Takes appropriately-sized mouthfuls	A. At kunne tage hensigtsmæssig størrelse mundfulde B. Tager mundfulde af passende størrelse C. Tager mundfulde i passende størrelse D. Tager passende mundfulde	E. Tager passende mundfulde CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner. *Kommentar: Upræcist hvad der er en passende mundfuld.* Ingen ændringer – afvent psykometrisk analyse.
9. Able to focus on meal	A. At kunne fokusere på måltidet B. Fokuserer på måltidet C. Er i stand til at fokusere på måltidet D. Kan fastholde opmærksomheden på måltidet	E. Kan fastholde opmærksomheden på måltidet CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
10. Demonstrates good judgment	A. At demonstrere god dømmekraft/ vurderingsevne B. Udviser god dømmekraft C. Udviser god dømmekraft D. Udviser god dømmekraft	E. Udviser god dømmekraft og adfærd. CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner. *Kommentar: Adfærd bør tilføjes item term, da det fremgår som en del af itembeskrivelse.* Adfærd tilføjet item term.
11. Tolerates physical effort of meal	A. At kunne klare fysisk udfordring ved indtagelse af måltidet/At tolerere fysisk indsats ved måltidet B. Tåler fysisk anstrengelse ved at spise C. Tolerer fysisk anstrengelse ved måltidet D. Tolererer måltidsaktivitetens krav	E. Kan udføre måltidet uden at udtrættes CVI=1/AD<0.65 for 3 indholdsvaliditetsdomæner. CVI=0.92 for ”adækvat item term”. *Kommentar: item undersøger udtrætning, så det er vel det item term skal beskrive.* Item term ændret
Liquid ingestion scale	A. Indtagelse af flydende føde skala B. Skala for indtagelse af væske C. Skala for indtagelse af drikke D. Skala for indtagelse af væsker	E. Skala for indtagelse af væsker
12. Seals lips on cup/glass	A. Tætner læber om kop/glas B. Lukker læberne om kop/glas C. Tilpasser og slutter læberne om kop/glas og holder læbeluk D. Tilpasser læbelukket til kop/glas	E. Tilpasser læbelukket til kop/glas CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner. *Kommentar: Hænger vel sammen med item 14 og 15.* Item bibeholdes - afvent psykometrisk analyse
13. Able to draw liquid from a standard straw	A. At kunne suge flydende væske gennem et almindeligt sugerør B. Suger væske med et almindeligt sugerør	E. Kan drikke med almindeligt sugerør CVI=1/AD<0.65 for 2 ind-

	C. Er i stand til at suge væske op via et almindeligt sugerør D. Kan drikke med almindeligt sugerør	holdsvaliditetsdomæner. CVI=0.92 for "klar scorebeskrivelse og CVI=0.85 for relevans. *Kommentar: Hvad er formålet med at kunne drikke med sugerør?* Item bibeholdt da brug af sugerør kan lette drikkefunktion ved perifer facialis parese + undersøger patientens oralmotoriske funktion.
14. Prevents leakage of liquid from cup/glass while drinking	A. At kunne drikke flydende væske af en kop/ et glas uden at spilde/ Spilder ikke væsken fra koppen/glasset imens der drikkes B. Undgår lækage af væske fra kop/glas, når der drikkes C. Forhindrer lækage af væske fra kop/glas under indtagelse af væske D. Drikker af kop eller glas uden at spilde. ("lækage" har flere betydninger – item term oversættes mhp. at reflektere itembeskrivelse).	E. Drikker af kop/glas uden der løber væske fra munden CVI=1/AD<0.65 for 3 indholdsvaliditetsdomæner. CVI=0.92 for relevans. *Kommentar: Item term beskriver ikke præcist hvornår der spildes.* Item term præciseret. *Kommentar: Slå sammen med item 15.* Item bibeholdt som separat item, da der er forskel på muskelfunktion i kinder/læber/tunge når væske "trækkes" ind i munden og når væsken kontrolleres inden i munden. - afvent psykometrisk analyse. *Kommentar: Uklart hvad en moderat mængde væske er i score 2.* Scorebeskrivelse præciseres og moderat mængde væske defineres som i item 15.
15. Prevents leakage of liquid from mouth before swallow	A. Spilder ikke væsken/At undgå at spilde flydende væske fra munden før man synker B. Undgår lækage af væske før synkning C. Forhindrer at væske spildes fra munden før synkning D. Holder væsken i munden uden at spilde (Som item 14).	E. Holder væsken i munden inden der synkes CVI=1/AD<0.65 for 2 indholdsvaliditetsdomæner. CVI=0.92 for "adækvat item term og CVI=0.92 for klar scorebeskrivelse. *Kommentar: synkning bør fremgå af item term.* Item term justeret. Uklar oversættelse og grammatiske fejl rettet i scorebeskrivelsen.
16. Able to take a sequence of sips	A. Kan tage flere slurke i træk B. Drikker med en sekvens af flere slurke	E. Kan drikke flere slurke ad gangen

	C. Er i stand til at tage flere på hinanden følgende slurke D. Kan drikke flere slurke ad gangen	CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
17. Demonstrates same voice quality after drinking	A. Demonstrerer/At have samme stemmeføring efter at have drukket B. Udviser samme stemmekvalitet efter drikning C. Kan demonstrere ensartet stemmekvalitet før og efter indtagelse af væske D. Har uændret stemmekvalitet efter at have drukket	E. Har uændret stemmekvalitet efter at have drukket CVI=1/AD<0.65 for 3 indholdsvaliditetsdomæner. CVI=0.85 for klar scorebeskrivelse. *Kommentar: hvordan kan man vurdere patientens stemmekvalitet, hvis man ikke må samtale med patient.* Under afsnittet "Tilrettelæggelse og forberedelse" tilføjes - Ergoterapeuten bør kun samtale med patienten med meget korte kommentarer for at opretholde den terapeutiske kontakt samt for at have mulighed for at vurdere patientens stemmekvalitet efter indtagelse af væske og fast føde *Kommentar: Hvad hvis patienten har afasi?* En sætning i beskrivelsen for score 1 er uddybet:" … or if he is unable to verbalize at the onset of the meal…." Uddybes til " eller hvis han fx pga. afasi ikke kan udtrykke sig verbalt ved måltidets begyndelse (overføres også til item 29).
18. Demonstrates clear airway after liquids	A. Demonstrerer/At demonstrere rene luftveje efter at have drukket B. Har rene luftveje efter væskeindtag C. Kan demonstrere frie luftveje efter indtagelse af væske D. Har rene luftveje efter at have drukket	E. Renser luftvejene, hvis der er behov efter indtagelse af væske. CVI=1/AD<0.65 for 3 indholdsvaliditetsdomæner. CVI=0.92 for klar scorebeskrivelse. *Kommentar: item term reflekterer ikke item- og scorebeskrivelsen. Der står, at det er patientens evne til at rense sit svælg efter penetration/aspiration.* Item term ændret (overføres også til item 30).
Solid ingestion scale	A. Indtagelse af fast føde skala B. Skala for indtagelse af fast føde C. Skala for indtagelse af mad D. Skala for indtagelse af fast føde	E. Skala for indtagelse af fast føde

19. Close upper lip on utensil	A. Tætner overlæben omkring redskabet B. Lukker overlæben om spiseredskabet C. Lukker overlæbe om bestik D. Former og slutter overlæben tæt til bestik	E. Former og slutter overlæben tæt til bestik CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
20. Prevents the loss of food from the mouth before swallowing	A. Spilder ikke maden ud af munden før den synkes B. Undgår tab af mad før synkning C. Forhindrer lækage af mad fra munden før synkning D. Holder maden i munden uden at spilde	E. Holder maden i munden inden der synkes CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner. *Kommentar: synkning bør fremgå af item term.* Item term justeret.
21. Use functional chewing pattern	A. At kunne bruge funktionelt tyggemønster B. Anvender funktionelt tyggemønstre C. Anvender et funktionelt tyggemønster D. Anvender et funktionelt tyggemønster	E. Anvender et funktionelt tyggemønster CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
22. Chewing appropriate to food item	A. Tygger maden hensigtsmæssigt/At tygge mad hensigtsmæssigt B. Tygger hensigtsmæssigt i forhold til fødeemner C. Tyggemetoden er i overensstemmelse med kosten D. Tygger hensigtsmæssigt i forhold til fødeemner	E. Tygger hensigtsmæssigt i forhold til fødeemner. CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner. *Kommentar: Spørgsmål til om det angivne antal på ca. 10 tyggesekvenser per mundfuld er korrekt.* Bibeholdt, da det også fremgår af teksten, at der er individuelle forskelle.
23. Positions bolus when chewing	A. Placering/Position af fødebolus når der tygges B. Anbringer/flytter bolus under tygning C. Bringer bolus i stilling under tygning D. Placerer bolus hensigtsmæssigt under tygning	E. Placerer bolus hensigtsmæssigt under tygning CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
24. Quantity of food remaining in mouth after swallow	A. Mængden af resterende mad i munden efter at patienten har sunket B. Mængder af mad, der resterer i munden efter synkning C. Mængden af madrester i munden efter synkning D. Mængden af madrester efter synkning	E. Mængden af madrester i munden efter synk CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
25. Location of food remaining in the mouth after swallow	A. Placering af madrester efter patienten har sunket/Placering af madresterne i munden efter at have sunket B. Placering af madrester efter synkning C. Madresteres placering i munden efter synkning D. Madresters placering i munden efter synkning	E. Madresters placering i munden efter synk CVI=1/AD<0.65 for 3 indholdsvaliditetsdomæner. CVI=0.92 for relevans. *Kommentar: kan evt. integreres i item 24.* Bibeholdt – - afvent psykometrisk analyse.
26. Swallow without extra effort	A. At synke uden anstrengelse B. Synker uden ekstra anstrengelse C. Synkning sker uden anstrengelse D. Synker uden anstrengelse	E. Synker uden anstrengelse CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
27. Swallows only once or	A. Kun at synke 1 eller 2 gange per mundfuld B. Synker kun 1 eller 2 gange per mundfuld	E. Synker kun 1 eller 2 gange per mundfuld

twice per mouthful	C. Synker kun 1 eller 2 gange per mundfuld D. Synker kun 1 eller 2 gange per mundfuld	CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
28. Maintains respiratory pattern throughout meal	A. At fastholde respiratorisk mønster under hele måltidet/Opretholder vejrtrækningsmønstret under hele måltidet B. Holder samme respiratoriske mønster under hele måltidet C. Opretholder normal/rytmisk åndedrætssekvens under måltidet D. Koordinerer åndedræt og spisning under måltidet. (Den direkte oversættelse gav uklar betydning af itembeskrivelse).	E. Koordinerer åndedræt og spisning under måltidet CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner.
29. Demonstrate same voice quality after eating	A. Demonstrerer/At have samme stemmeføring efter at have spist B. Udviser samme stemmekvalitet efter som før måltidet C. Kan demonstrere ensartet stemmekvalitet før og efter indtagels af fast kost D. Udviser uændret stemmekvalitet efter at have spist	E. Har uændret stemmekvalitet efter at have spist CVI=1/AD<0.65 for alle indholdsvaliditetsdomæner. *Samme kommentarer som item 17.*
30. Demonstrates clear airway after solids	A. At demonstrere/at have rene luftveje efter at have spist B. Har rene luftveje efter indtagelse af fast føde C. Kan demonstrerer frie luftveje efter indtagels af fast kost D. Har rene luftveje efter at have spist	E. Renser luftvejene, hvis der er behov efter indtagelse af fast føde CVI=1/AD<0.65 for 2 indholdsvaliditetsdomæner. CVI=0.92 for "henholdsvis klar item- og klar scorebeskrivelse. *Samme kommentarer som item 18.*
Texture management - solids	A. Håndtering af konsistens skala – fast føde B. Skala for håndtering af konsistens - fast føde C. Skala for konsistens håndtering – fast konsistens D. Skala for håndtering af fast konsistens	E. Skala for håndtering af fast konsistens
Texture management - liquids	A. Håndtering af konsistens skala – væske B. Skala for håndtering af konsistens – væsker C. Skala for konsistens håndtering – væske konsistens D. Skala for håndtering af væske konsistens	E. Skala for håndtering af væske konsistens
Ekspertvurderingen resulterede i at items i begge skalaer opnåede CVI=1/AD<0.65 for 3 indholdsvaliditetsdomæner og CVI=0.92 for klar itembeskrivelse. *Kommentar: er det sikkerhed når patienten synker de forskellige konsistenser eller er det patientens kognitive funktion? - der er måske mange formål- det scores jo også under skalaerne for indtagelse af væske og fast føde.* -Afvent psykometrisk analyse før der ændres. Hvis der ændres nu så vil items og skalaerne blive meget forskellig fra original versionen. Nuværende grundlag for ændring er for spinkelt til at få godkendt så omfangsrige ændringer fra CAOT, der har Copyright på MISA. - Scorebeskrivelserne opstilles med samme layout som for øvrige items i MISA (er ikke tilfældet i originalversionen). Alle konsistenstyper (item 31-43), deres beskrivelser og eksempler på fødevarer opstilles i et skema (er ikke tilfældet i originalversionen). Er godkendt af CAOT og Heather.		
31. Capable of eating heterogeneous	A, B, C. Kan spise heterogen konsistens D. Kan spise heterogent/blandet konsistens (For eksemplerne på konsistensen udelades	E. Kan spise heterogent/blandet konsistens

	Shepherd's pie og der tilføjes millionbøf med kartoffelmos, boller i ris og karry, rugbrød med skivepålæg/ost).	
32. Capable of eating fibrous solids	A. Kan spise trævlet fast føde B. Kan spise fiberholdigt fast føde C. Kan spise fibrøs fast konsistens D. Kan spise trævlet konsistens	E. Kan spise trævlet konsistens
33. Capable of eating hard solids.	A, B. Kan spise hård fast føde C. Kan spise hård fast konsistens D. Kan spise hård konsistens (For eksempler på konsistensen tilføjes tvebakker, kammerjunker og rugbrød uden kærner).	E. Kan spise hård konsistens
34. Capable of eating minced/granular solids	A. Kan spise finthakket/kornet fast føde B. Kan spise hakket/kornet fast føde C. Kan spise hakket/granuleret fast konsistens D. Kan spise hakket/granuleret konsistens (For eksempler på konsistensen tilføjes bulgur, solsikke- og pinjekerner).	E. Kan spise hakket/ granuleret konsistens
35. Capable of eating sticky solids	A. Kan spise klæbrig fast føde B. Kan spise klistret fast føde C. Kan spise klæbrig fast konsistens D. Kan spise klæbrig konsistens (For eksempler på konsistensen tilføjes chokolade, Nutella og leverpostej).	E. Kan spise klæbrig konsistens
36. Capable of eating soft solids	A. Kan spise blød konsistens B. Kan spise blød fast føde C. Kan spise blød fast konsistens D. Kan spise blød konsistens	E. Kan spise blød konsistens
37. Capable of eating puree	A, B, C, D. Kan spise puré	E. Kan spise puré
38. Capable of eating pudding	A, B. Kan spise budding C. Kan spise budding konsistens (fast) D. Kan spise budding	E. Kan spise budding
39. Capable of drinking water	A, B, C,D. Kan drikke vand	E. Kan drikke vand
40. Capable of drinking thin juices	A. Kan drikke tynd juice B. Kan drikke tynd saft C. Kan drikke tynd væske/juice D. Kan drikke tynd væske (Kategorien inkluderer kaffe/the, mælk og sorbet is – derfor vil "juice" være misvissende)	E. Kan drikke tynd væske
41. Capable of drinking nectar consistency liquids	A. Kan drikke væske med konsistens af nektar B. Kan drikke nektar-lignende væske C. Kan drikke nektar konsistens D. Kan drikke nektar konsistens	E. Kan drikke nektar konsistens
42. Capable of drinking honey consistency liquids	A. Kan drikke væske med konsistens af honning B. Kan drikke honninglignende væske C. Kan drikke honning konsistens D. Kan drikke honning konsistens	E. Kan drikke honning konsistens
43. Capable of drinking pudding consistency liq-	A. Kan drikke væske med konsistens af budding B. Kan drikke budding-lignende væske C. Kan drikke budding konsistens (Væske)	E. Kan drikke budding konsistens

uids	D. Kan drikke budding konsistens	
Scorekategorierne 1 til 3 er defineret specifikt for hvert item i instruktionsmanualen og er forkortet på registreringsskemaet for 18 items. For 25 items fremgår scorerne som identiske kategorier:		
Original version	Første oversættelser: 3 oversættere (A, B og C) Konsensusoversættelse: Review gruppe og semantisk kontrol (D)	Endelig oversættelse efter ekspertpanel vurdering (E)
1= newer or rarely 2= Sometimes 3= Always or almost always	A. (1= Aldrig eller sjældent; 2= Nogle gange; 3= Altid eller næsten altid) B. (1= På intet tidspunkt eller sjældent; 2= Af og til; 3=Altid eller næsten altid) C. (1=Aldrig eller sjældent; 2= Af og til; 3= Altid eller næsten altid) D. (1= På intet tidspunkt eller sjældent; 2= Indimellem; 3= Altid eller næsten altid)	E. 1= På intet tidspunkt eller sjældent 2= Indimellem 3= Altid eller næsten altid
Note: Oversættelse B blev brugt som ”grundstamme” for oversættelsen af instruktionsmanualen og registreringsarket. Elementer fra de to andre versioner blev integreret hvor det var relevant. Review gruppen vurderede oversættelse B til at være den mest præcise og sprogligt flydende oversættelse. Oversætter B blev kontaktet ved behov.		
Instruktionsmanual og registreringsark		
Note: I afsnittet om registrering og scoring var der en fejl i original versionen mht. hvordan den procentvise skalascore udregnes. Dette er korrigeret i den danske version. Alle øvrige afsnit blev godkendt af ekspertpanelet med CVI=1 og AD< 0.65.		

Appendix B - The McGill Ingestive Skills Assessment (Danish version).

Følgende registreringsskema og uddrag af instruktionsmanualen (Referenceramme, Anvendelse af MISA og MISA-Score) er projektudgaven og må ikke kopieres.

McGill Ingestive Skills Assessment

Patient navn:_____	**Afdeling:**_____	**Diagnose(r):**
CPR:_____	**Stuenummer:**_____	
Ergoterapeut:	**Dato:**	

Hjælpemidler:

Skala	**# items**	**Patientscore**	**Maksimum**	**% score** (Patientscore - # items) x 100 (maksimum - # items)
Siddestilling	4		12	
Spise- og drikkefærdigheder	7		21	
Indtagelse af væsker	7		21	
Indtagelse af fast føde	12		36	
Håndtering af fast konsistens	8		24	
Håndtering af væske konsistens	5		15	
TOTAL	43		129	

Siddestilling	**1 point**	**2 point**	**3 point**
Opretholder symmetrisk kropsstilling	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Opretholder passende hovedstilling I forhold til at spise og drikke	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Opretholder 90 graders hoftefleksion	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Opretholder postural stabilitet I truncus	Behøver vedvarende støtte med armen	Behøver lejligheds-vis støtte	Behøver ikke støtte med armene for at opretholde siddestillin-gen

Spise- og drikkefærdigheder	1 point	2 point	3 point
Kan tage funktionelt fat om bestik/fødeemne og føre det til munden	På intet tidspunkt eller sjældent, eller spiser ikke selv	Indimellem	Altid eller næsten altid
Kan tage funktionelt fat om kop/glas og føre det til munden	På intet tidspunkt eller sjældent, eller spiser ikke selv	Indimellem	Altid eller næsten altid
Vælger hensigtsmæssigt bestik i forhold til fødeemne	På intet tidspunkt eller sjældent, eller spiser ikke selv	Indimellem	Altid eller næsten altid
Tager passende mundfulde	På intet tidspunkt eller sjældent, eller spiser ikke selv	Indimellem	Altid eller næsten altid
Kan fastholde opmærksomheden på måltidet	Kan ikke fastholde opmærksomheden	Distraheres lejlighedsvis	Kan forblive opmærksom
Udviser god dømmekraft og adfærd	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan udføre måltidet uden at udtrættes	Er udtrættet ved begyndelsen af måltidet	Udtrættes undervejs i måltidet	Der observeres ingen udtrætning

Indtagelse af væsker	1 point	2 point	3 point
Tilpasser læbelukket til kop/glas	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan drikke med almindeligt sugerør	Suger ikke, eller suger ikke kraftigt nok	Drikker med besvær	Drikker uden besvær
Drikker af kop/glas uden der løber væske fra munden	Der løber store mængder	Der løber små til moderate mængder	Der løber ikke væske fra patientens mund
Holder væsken i munden inden der synkes	Der løber store mængder	Der løber små til moderate mængder	Der løber ikke væske fra patientens mund
Kan drikke flere slurke ad gangen	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Har uændret stemmekvalitet efter at have drukket	Fuldstændigt tab af stemme, eller stemmen bliver våd, hæs eller gurglende efter indtag af små mængder væske, eller taler ikke trods tilskyndelse	Stemmen bliver våd, hæs eller gurglende efter indtag af store mængder væske	Ingen ændring af stemmen efter indtag af væske
Renser luftvejene, hvis der er behov efter indtagelse af væske	Renser ikke svælget ved behov, eller har ineffektiv rensning af svælget	Renser svælget effektivt	Behøver ikke at rense svælget under måltidet

Indtagelse af fast føde	1 point	2 point	3 point
Former og slutter overlæben tæt til bestik	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Holder maden i munden inden der synkes	Vedvarende tab, eller taber indimellem store mængder	Tab af små mængder	Intet tab
Anvender et funktionelt tyggemønster	Ingen tyggeforsøg, eller sutter kun	Vertikale bevægelser - gumlen	Normalt roterende tyggemønster
Tygger hensigtsmæssigt i forhold til fødeemner	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Placerer bolus hensigtsmæssigt under tygning	Mod den hårde gane, eller former ikke en sammenhængende bolus, eller tygger ikke	På fortænderne, eller indimellem på kindtænderne	Altid på kindtænderne
Mængder af madrester i munden efter synk	Der resterer mere end halvdelen af bolus	Der resterer mindre end halvdelen af bolus	Ingen madrester
Madresters placering i munden efter synk	Klæber mod den hårde gane, eller i kindhulerne	Omkring tungen eller på tænderne	Ingen madrester
Synker uden anstrengelse	På intet tidspunkt eller sjældent	Indimellem, eller kun ved visse typer af faste konsistenser	Altid eller næsten altid
Synker kun 1 eller 2 gange per mundfuld	På intet tidspunkt eller sjældent	Indimellem, eller kun ved visse typer af faste konsistenser	Altid eller næsten altid
Koordinerer åndedræt og spisning under måltidet	Lejlighedsvis svære vanskeligheder, eller mindre vanskeligheder under hele måltidet	Lejlighedsvis mindre vanskeligheder	Ingen vanskeligheder
Har uændret stemmekvalitet efter at have spist	Fuldstændigt tab af stemme, eller stemmen bliver våd, hæs eller gurglende efter indtag af små mængder fast føde, eller taler ikke trods tilskyndelse	Stemmen bliver våd, hæs eller gurglende efter indtag af store mængder fast føde	Ingen ændring i stemmen efter indtag af fast føde
Renser luftvejene, hvis der er behov efter indtagelse af fast føde	Renser ikke svælget ved behov, eller har ineffektiv rensning af svælget	Renser svælget effektivt	Behøver ikke at rense svælget

Håndtering af fast konsistens	1 point	2 point	3 point
Kan spise heterogen/blandet konsistens	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan spise trævlet konsistens	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan spise hård konsistens	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan spise hakket/ granuleret konsistens	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan spise klæbrig konsistens	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan spise blød konsistens	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan spise puré	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan spise budding	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid

Håndtering af væske konsistens	1 point	2 point	3 point
Kan drikke vand	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan drikke tynd væske	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan drikke nektar konsistens	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan drikke honning konsistens	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid
Kan drikke budding konsistens	På intet tidspunkt eller sjældent	Indimellem	Altid eller næsten altid

Referenceramme

"The McGill Ingestive Skills Assessment" (MISA) er et undersøgelsesredskab, der fokuserer på patientens evne til sikkert og selvstændigt at indtage flere forskellige faste og flydende konsistenser. Opfattelsen af spise- og drikkeproblemer har udviklet sig til at inkludere områder udover synkningens fysiologi og dens relaterede problemer (dvs. fejlsynkning og dysfagi). Undersøgelser af evnen til at spise, drikke og synke omfatter nu også det, der sker før maden når munden og som påvirker kvaliteten af måltidet. Den normale synkeproces inkluderer: fysiologiske faktorer som sult, sansefunktion og perception fx, at maden bliver set og duftet, kognition, motoriske færdigheder samt omgivelserne omkring måltidsaktiviteten (en samling eller "konstruktion" af færdigheder, der på engelsk benævnes "Ingestion"[1] (Leopold og Kagel, 1997)). Denne bredere definition af, hvad der udgør et spise-, drikke- og synkeproblem blev anvendt i teoretiseringen og udviklingen af MISA (Lambert, Gisel, Wood-Dauphinee og Groher, 2003). Det nye aspekt som MISA bidrager med er, at der lægges vægt på funktionel undersøgelse frem for fysiologisk undersøgelse. MISA undersøger evnen til sikkert og selvstændigt at indtage et bredt udvalg af mad og drikkevarer samt evnen til at deltage i en måltidsrutine hos personer med neurologiske dysfunktioner.

[1] O.A. Den danske oversættelse af "Ingestion", er ernæringsmæssig indtagelse. Denne direkte oversættelse synes ikke at være dækkende for definitionen i Leopold og Kagel (1997). Derfor oversættes "Ingestion" ikke direkte i dette afsnit.

Anvendelsen af MISA

MISA består af 43 items inddelt i 6 skalaer: siddestilling ved måltider (4 items), færdigheder i forbindelse med at spise og drikke selv (7 items), oral motoriske færdigheder ved indtagelse af væsker (7 items) og af fast føde (12 items), samt konsistenshåndtering (8 items for faste og 5 items for væske konsistenser). Items scores på en 3-points ordinal skala, hvor 3 indikerer sikker og selvstændig udførelse under måltidet, 2 indikerer ineffektiv og/eller mangelfuld udførelse under måltidet, og 1 indikerer ingen sikker og selvstændig udførelse under måltidet. Scorerne opsummeres til en skalascore indenfor hver skala og en totalscore for hele MISA-undersøgelsen.

Hensigten med MISA

MISA er et undersøgelsesredskab, der fokuserer på patientens evne til selvstændigt og sikkert at indtage et bredt udvalg af faste konsistenser og væske konsistenser. Formålet med MISA er at bibringe ergoterapeuten information vedrørende en persons aktuelle daglige funktion, hvilket derudover også kan bruges til at supplere den information, der opnås via en neurologisk og fysiologisk undersøgelse. MISA kvantificerer sikkert og pålideligt en persons spise-, drikke- og synkefærdigheder, uden at det kræver fuldførelsen af specifikke tests. MISA tilvejebringer kvantitative indikatorer på, hvordan patienten fungerer. MISA er et diagnostisk undersøgelsesredskab, der tildeler en numerisk score til patientens funktionsevne og identificerer områder, hvor færdighederne er svækket, hvilket kan guide ergoterapeuten i rehabiliteringsplanlægningen (Lambert et al., 2003).

MISA er designet til at blive anvendt i vurderingen af og behandlingsplanlægningen til voksne og ældre personer med færdighedstab i forhold til at spise, drikke og synke; et færdighedstab som er af neurologisk snarere end kirurgisk eller neoplastisk oprindelse. MISA er testet på personer med Demens, Parkinson, Apopleksi, Multipel Scle-

rose, Amyotrofisk Lateral Sclerose samt personer i en fremskreden alder. MISA er ikke testet på personer med laryngektomi, glossektomi eller hoved- og halskræft.

Størstedelen af MISA's testegenskaber (*reliabilitet og validitet*) er etableret på personer under akut indlæggelse, på personer med behov for langvarig pleje samt på personer på genoptræningscentre. Informationerne om prædikativ validitet blev udelukkende etableret på personer med behov for langvarig pleje. Selvom MISA kan anvendes til at evaluere personer fra enhver af de nævnte områder, så er informationen om prædikativ validitet mest brugbar i forhold til personer med behov for langvarig pleje.

Forudsætning for at anvende MISA

MISA administreres via observation af et måltid. Inden MISA anvendes, bør ergoterapeuten være fortrolig med alle items og deres scoring. Det forudsættes derfor, at ergoterapeuten har sat sig grundigt ind i denne instruktionsmanual, og har fået indgående kendskab til items før MISA anvendes. Det anbefales, at ergoterapeuten observerer og MISA-scorer mindst 5 patienter inden MISA bruges til vurdering og behandlingsplanlægning. Dette er nødvendigt for at kunne anvende MISA korrekt. Indgående kendskab til items betyder, at ergoterapeuten benytter tiden til at observere og vurdere patienten under måltidet, i stedet for at prøve at finde items på undersøgelsesskemaet og huske hvordan de skal scores.

Testmåltid

MISA-undersøgelsen nødvendiggør ikke, at patienten gives et bestemt måltid. På denne måde kan der tages hensyn til individuelle madpræferencer eller diætrestriktioner. Det anbefales dog på det kraftigste, at der tilpasses et specifikt testmåltid indenfor den enkelte afdeling/institution. Dette kan være med til at lette planlægningen af en MISA-undersøgelse samt maksimere ensartetheden i testomstændighederne fra en undersøgelse til en anden.

Testmåltidet bør indeholde alle de konsistenser af mad og drikke, der er beskrevet i MISA. Der bør serveres små portioner af hver konsistenstype. Dette kan nødvendiggøre, at patienten præsenteres for et større måltid end sædvanligt. Bemærk dog, at patienten ikke mister points, hvis han ikke spiser maden op. Patienten behøver kun at spise og drikke så meget af hver enkelt konsistens, at ergoterapeuten har mulighed for at vurdere patientens sikkerhed ved indtagelse af konsistensen samt kan evaluere de items, der er specifikke for oralmotoriske færdigheder.

Hvis det allerede er vurderet, at patienten ikke kan indtage visse konsistenser sikkert, så kan disse udelukkes fra måltidet. Patienten tildeles da den lavest mulige score for det pågældende item på skalaerne for håndtering af konsistenser.

Et testmåltid kan fx indeholde små portioner som i det følgende skema.

Fødeemne	**Konsistenstyper**
Hønsekødssuppe med grøntsager	Heterogen / blandet konsistens
Steg af okse-, kalve- eller svinekød	Trævlet fast konsistens
Tilberedte gulerødder	Blød fast konsistens
Ris	Hakket / granuleret fast konsistens
Kartoffelmos	Klæbrig føde fast konsistens
Småkager	Hård fast konsistens
Budding	Budding fast konsistens
Fin æblemos	Puré
Vand	Vand
Te eller kaffe	Tynd væske
Sveske juice	Væske af nektar konsistens
Fortykket væske	Væske af honning konsistens
Fortykket væske	Væske af budding konsistens

Tilrettelæggelse og forberedelse

For at sikre sig, at alle de nødvendige konsistenstyper og bestik er til stede, bør ergoterapeuten ankomme til patientens spisested i god tid før måltidets begyndelse. Ergoterapeuten bør også sikre sig, at patienten har sit sædvanlige vågenhedsniveau og ikke er nævneværdig påvirket af faktorer som ændring i medicinering, stressende medicinske undersøgelser eller et særlig travlt genoptræningsprogram. Ergoterapeuten sidder formentlig bedst i en afstand af ca. 2 meter fra patienten og placeret således, at patientens ansigt ses i halvprofil. Fra den vinkel, har ergoterapeuten mulighed for at se ind i patientens mund fra tid til anden, uden at virke forstyrrende. Ergoterapeuten bør kun samtale med patienten med meget korte kommentarer for at opretholde den terapeutiske kontakt samt for at have mulighed for at vurdere patientens stemmekvalitet efter indtagelse af væske og fast føde. Med mindre andet er angivet i instruktionsmanualens skala- og item beskrivelser, bør ergoterapeuten ikke intervenere i forhold til patientens siddestilling under måltidet eller i forhold til patientens udførelse, når han spiser og drikker.

For at sikre scoring af de forskellige items, kan patienten tilskyndes til at prøve forskellige fødeemner og væsker. Hvis patienten normalt assisteres af en plejer, portør, familiemedlem eller en frivillig, så bør denne person yde den sædvanlige assistance under MISA-undersøgelsen. **Undersøgelsen skal afbrydes, hvis måltidssituationen vurderes til at være farlig for patienten, hjælperen eller ergoterapeuten.**

Registrering og scoring

Brugere af MISA har fundet det mest effektivt at observere omkring halvdelen af måltidet, før undersøgelsesskemaet udfyldes. Under hver sektion i undersøgelsesskemaet er der plads til at gøre notater undervejs i observationen. Hvis der er items, der ikke er observeret, så scores patienten med den lavest mulige score. Hvis patientens udførelse er på et niveau mellem to scorebeskrivelser, eller hvis patientens udførelse er på ét niveau på ét tidspunkt og på et lavere niveau på ét andet tidspunkt, så gives der altid den laveste score. Hvert items' score markeres tydeligt med en kuglepen, highlighter eller blyant. Hvilken markering, der bruges er ikke vigtig, men vær derimod omhyggelig med, at alle items er markeret og at der ikke kan herske tvivl om, hvilken score, der er markeret.

Skalascoren udregnes som summen af items indenfor den enkelte skala. For at lette tolkningen af resultaterne kan den procentvise score for hver skala udregnes (den opnåede score minus antal items i skalaen, divideret med maksimumscoren for skalaen minus antal items i skalaen og ganget med 100). En totalscore og den procentvise totalscore kan udregnes på samme måde. I opsummeringsfeltet på undersøgelsesskemaets forside, er der plads til at notere de hjælpemidler som patienten bruger under måltidet. Noter både store hjælpemidler som kørestol eller lejringshjælpemidler og/eller små hjælpemidler som specielbestik og specielle kopper/glas.

Læs omhyggelig de følgende sider med beskrivelser af hver MISA-skala og scoringen af hvert item indenfor hver skala. For læsevenlighedens skyld, bruges stedordet "han" for både mandlige og kvindelige patienter.

MISA-Score

Skala for siddestilling

Formålet med skalaen er at vurdere patientens evne til at opretholde en hensigtsmæssig siddestilling i forhold til at spise og drikke. Ergoterapeuten bør som hovedregel ikke ændre patientens siddestilling under måltidet; den opgave bør være op til den sædvanlige plejer eller patienten selv. Hvis ergoterapeuten er eneste sundhedsprofessionelle tilstede og patienten ikke kan opretholde en hensigtsmæssig siddestilling, så kan ergoterapeuten intervenere. Dette skal dog afspejles i de givne scorer.

Opretholder symmetrisk kropsstilling

Under dette item, observeres om patienten læner sig mod højre eller venstre under måltidet. Patienten bør indtage en symmetrisk kropsstilling, når han tager en mundfuld eller synker.

Score 1 hvis patienten, hele tiden eller næsten hele tiden, læner sig til én side under måltidet og ikke retter sig op, selv når han tager en ny mundfuld eller synker.

Score 2 hvis patienten læner sig til en side noget af tiden under måltidet. Måske læner patienten sig indimellem til siden, men retter sig op, eller måske mister patienten kontrollen over sin stilling og genvinder ikke symmetrisk kropsstilling under resten af måltidet.

Score 3 hvis patienten kan opretholde en symmetrisk kropsstilling, eller hvis han kun ganske kortvarigt indtager en asymmetrisk stilling, men retter sig op, når han tager en mundfuld eller synker. Hvis ikke han gør det, scores 2. Nogle flytter vanemæssigt kropsvægten fra side til side; hvis de kommer tilbage til midtlinjen for at tage en mundfuld mad eller drikke, scores 3.

Opretholder passende hovedstilling i forhold til at spise og drikke

Under dette item, observeres patientens evne til at holde hovedet i en neutral stilling; dvs., at hovedet er i kroppens midtlinje og let flekteret med hagen mod brystet. Under indtagelse af mad og drikke, bør hoved og nakke ikke være ekstenderet, roteret eller lateral flekteret (medmindre dette er ordineret i forvejen som kompensation for et synkeproblem; og så scores 3).

Score 1 hvis patienten ikke har hovedkontrol og har behov for støtte med hjælpemiddel eller kæbekontrolgreb, eller hvis han på intet tidspunkt kan holde hovedet i en stabil neutral stilling.

Score 2 hvis patienten kan indtage og opretholde en neutral hovedstilling nogle få sekunder eller minutter ad gangen, men ikke har hovedkontrol under størstedelen af måltidet, eller hvis patienten ikke har hovedkontrol i en længere periode under måltidet.

Score 3 hvis patienten opretholder en neutral og stabil hovedstilling under hele måltidet. Nogle vil måske kigge sig omkring i rummet, eller bøje og strække sig med vilje for at løsne stivhed; hvis disse bevægelser er viljestyrrede, scores 3. Hvis der forekommer ufrivillige bevægelser samtidig med de viljestyrrede bevægelser, scores 2.

Opretholder 90 graders hoftefleksion

Under dette item, observeres om patienten glider frem i stolen og dermed åbner hoftevinklen. I den siddende stilling ved indtagelse af mad og drikke, bør bækkenet være fremadkippet således, at hoften er flekteret svt. 90 grader. Hvis der bruges et lejringshjælpemiddel til at opretholde en hensigtsmæssig siddestilling, så trækker det ikke ned i scoren medmindre patienten alligevel glider frem. Noter altid, hvilket lejringshjælpemiddel patienten bruger, også selvom han opretholder 90 graders hoftefleksion.

Score 1 hvis patienten glider så meget frem, at gentagne korrektioner af siddestil-

lingen er nødvendig under måltidet for at forhindre, at han glider ned fra stolen.

Score 2 hvis patienten gradvist glider frem under måltidet, således at én korrektion af siddestillingen er nødvendig.

Score 3 hvis patienten glider en lille smule frem, men ikke har behov for korrektion af siddestillingen, eller hvis patienten slet ikke glider frem under måltidet.

Opretholder postural stabilitet i truncus

Under dette item, observeres patientens generelle posturale stabilitet samt om han har tilstrækkelig styrke til at stabilisere i truncus.

Score 1 hvis patienten konstant støtter sig med armen, eller har behov for hjælpemidler til at opretholde stabilitet i truncus, eller har behov for at være i en halvsiddende stilling, fx i en bagud kippet stol.

Score 2 hvis patienten indimellem støtter sig med armen under måltidet, eller hvis han begynder at støtte sig med armen undervejs i måltidet, eller hvis han bruger armen til at korrigere sin siddestilling.

Score 3 hvis patienten ikke bruger armen til at stabilisere sig og ikke bruger armen, når han flytter sin kropsvægt eller skifter stilling. Nogle vil hvile armene på bordet eller stolearmlænene mens de spiser, og scores 3. Hvis patienten holder en arm således, men indimellem bruger armen til støtte eller til at korrigere sin siddestilling, scores 2.

Skala for spise- og drikkefærdigheder

Formålet med skalaen er at vurdere patientens evne til at klare måltidsaktiviteten selvstændigt og funktionelt. Skalaen berører forskellige aspekter af et måltid, herunder at spise og drikke selv, adfærd og dømmekraft.

Kan tage funktionelt fat om bestik/fødeemne og føre det til munden

Under dette item observeres, hvordan patienten holder bestikket. Måden, hvorpå patienten tager fat om bestik er ikke vigtig. Til gengæld, bør bestikket holdes således, at det er muligt at tage mad op fra tallerkenen samt at forhindre, at mad spildes før bestikket når munden. Hvis patienten normalt bruger hånden til visse fødeemner, fx et stykke franskbrød med ost eller marmelade, så observeres dette også. Men vær sikker på, at testmåltidet også indeholder fødeemner, der kræver brug af bestik. Observer i forhold til, hvordan patienten udfører begge metoder og giv en score på den metode, der volder patienten de største problemer. Notér hvilken af metoderne der henholdsvis var sværest og lettest.

Score 1 hvis patienten mades af en hjælper, eller hvis han sjældent tager funktionelt fat om bestikket/fødeemnet. Hvis patienten guides fysisk, scores1.

Score 2 hvis patienten spiser selvstændigt og han tager funktionelt fat om bestikket/fødeemnet i dele af måltidet. Patienten har måske et funktionelt greb og skifter senere til et ikke-funktionelt greb (eller omvendt), eller patienten holder funktionelt om bestikket/fødeemnet, når han instrueres af en hjælper eller får bestikket/fødeemnet placeret i hånden. Maden spildes måske fra bestikket eller fødeemnet tabes mere end to gange.

Score 3 hvis patienten spiser selvstændigt og holder funktionelt om bestikket/fødeemnet under hele måltidet, kun spilder mad fra bestikket eller taber fødeemnet én eller to gange under måltidet og ikke har behov for hjælp til at opnå et funktionelt greb.

Kan tage funktionelt fat om kop/glas og føre det til munden

Under dette item, observeres patienten, når han drikker af en kop og/eller af et glas. Koppen/glasset bør ikke fyldes mere end 1 cm fra kanten. Måden, hvorpå patienten tager fat om kop/glas er ikke vigtig. Til gengæld bør koppen/glasset holdes så spild af

væsken forhindres inden munden nås. Hvis der er hank på koppen, behøver patienten ikke at bruge den.

Score 1 hvis patienten mades af en hjælper, eller hvis han sjældent tager funktionelt fat om kop/glas. Hvis patienten drikker al væske selvstændigt med en ske, eller hvis patienten guides fysisk, scores 1.

Score 2 hvis patienten drikker selvstændigt og tager funktionelt fat om kop/glas i dele af måltidet, men der spildes mere end to gange. Hvis patienten selv drikker, men har behov for at kop/glas placeres i hånden, scores 2.

Score 3 hvis patienten bruger et funktionelt greb under hele måltidet og kun spilder væske fra kop/glas én eller to gange, når der drikkes.

Vælger hensigtsmæssigt bestik i forhold til fødeemne

Under dette item observeres, hvilket bestik patienten vælger for hvert fødeemne. For at afgøre om det valgte bestik er hensigtsmæssigt, bør både fødeemnets konsistens og patientens sikkerhed tages i betragtning. Hvis patienten har nedsat finmotorisk kontrol og er i risiko for at beskadige læberne, så er en gaffel muligvis ikke et hensigtsmæssigt bestik. Det er på intet tidspunkt hensigtsmæssigt at spise af en kniv. Hvis patienten normalt bruger hånden til visse fødeemner, fx et stykke franskbrød med ost eller marmelade, scorer det ikke ned. Men vær sikker på at testmåltidet også indeholder fødeemner, der kræver brugen af bestik.

Score 1 hvis patienten mades af en hjælper, hvis en anden person vælger bestikket, eller hvis patienten næsten aldrig vælger hensigtsmæssigt bestik.

Score 2 hvis patienten indimellem kan vælge hensigtsmæssigt bestik.

Score 3 hvis patienten altid eller næsten altid vælger hensigtsmæssigt bestik.

Tager passende mundfulde

Under dette item, observeres den mængde mad patienten tager i munden. Bolus bør være stor nok til, at den kan tygges, men ikke så stor, at kinderne fyldes helt ud. En teskefuld puré med top eller en gaffelfuld er passende for de fleste. Der kan dog være undtagelser. Fx kan en patient med en lille mundhule have behov for at tage mindre mundfulde, og en patient med nedsat sensibilitet kan have behov for at komme mere mad i munden på én gang, for at kunne mærke hvor maden er. Brug din dømmekraft og viden om patientens situation til at vurdere, hvad der er passende for den enkelte patient.

Score 1 hvis patienten altid eller næsten altid putter for meget eller for lidt mad i munden, eller hvis han mades af en hjælper.

Score 2 hvis patienten indimellem putter for meget eller for lidt mad i munden.

Score 3 hvis patienten konsekvent tager passende mundfulde.

Kan fastholde opmærksomheden på måltidet

Under dette item, observeres patientens evne til at fastholde opmærksomheden på måltidet. At spise er en social aktivitet, og det er passende at interagere med andre tilstedeværende under måltidet. Distraherbarhed bliver et problem, når patienten afledes af hændelser, der ikke er relateret til ham selv og/eller måltidsaktiviteten (fx, en plejer, der går gennem rummet, en personsøger, der ringer, fnug på tøjet m.m.), eller hvis patienten bliver overstimuleret af samtale. I begyndelsen af MISA-undersøgelsen, informeres patienten om, at du er der for at observere hans formåen til at spise, drikke og synke, og at han derfor ikke skal være opmærksom på dig. En vis interaktion er dog uundgåelig trods denne påmindelse, men den bør være tilstrækkeligt til at afværge overdreven snak.

Score 1 hvis patienten distraheres og ikke flytter opmærksomheden tilbage og fortsætter måltidet, eller hvis han hjælpes til at flytte opmærksomheden tilbage og fortsætte måltidet.

Score 2 hvis patienten distraheres, men selv flytter opmærksomheden tilbage og

fortsætter måltidet.

Score 3 hvis patienten ikke distraheres fra måltidet, og hvis han kun samtaler i et socialt passende omfang. Hvis patienten samtaler overdrevent, scores 2.

Udviser god dømmekraft og adfærd

Under dette item, observeres patientens dømmekraft; dvs. de beslutninger, der træffes under et måltid for at sikre sin personlige sikkerhed og for at opretholde en passende adfærd under måltidet. Det indbefatter at synke før den næste mundfuld tages, ikke at tale med mad i munden, og ikke at spytte, kaste eller stjæle mad. Notér enhver adfærd, der tyder på ringe dømmekraft.

Score 1 hvis patienten mades og vil acceptere den næste mundfuld uden at synke den foregående, eller hvis patienten spiser og drikker selvstændigt, men udviser ringe dømmekraft under hele måltidet.

Score 2 hvis patienten udviser ringe dømmekraft gentagne gange under måltidet.

Score 3 hvis patienten udviser god dømmekraft i løbet af hele måltidet, eller hvis patienten mades og afviser at få mere mad før han har tygget af munden.

Kan udføre måltidet uden at udtrættes

Under dette item, observeres patientens energiniveau periodevis under hele måltidet. Patienten bør være ved fuld bevidsthed under hele måltidet. Læg mærke til tegn på udtrætning, såsom at patienten gaber, har lukkede eller udtryksløse øjne, falder i søvn, får nedsat spise- og drikketempo, eller bruger længere tid på at bearbejde og synke hver mundfuld.

Score 1 hvis patienten er træt ved begyndelsen af måltidet.

Score 2 hvis patienten udtrættes undervejs i måltidet.

Score 3 hvis patienten indtager hele måltidet uden at udtrættes.

Skala for indtagelse af væsker

Formålet med skalaen er at vurdere patientens evne til at indtage almindelige og fortykkede væsker. Skalaen skal scores, uanset om patienten drikker af kop, glas eller bruger en ske. Skalaen omhandler de forskellige motoriske færdigheder, der er forbundet med at drikke samt de observerbare tegn på faryngeal dysfunktion og svækket luftvejsbeskyttelse.

Tilpasser læbelukket til kop/glas

Under dette item, observeres patientens evne til at tilpasse læbelukket til koppen/glasset, når han drikker; vær især opmærksom på underlæben. Når patienten drikker, bør kop/glas hvile på underlæben og overlæben bør slutte ned mod koppens/glassets inderside. Det er dog ikke alle, der former overlæben så den slutter helt tæt til koppens/glassets overflade; det regnes for helt normalt.

Score 1 hvis patienten næsten aldrig tilpasser læbelukket til kop/glas. Dette kan ses ved, at patienten stabiliserer koppen/glasset på undertænderne, og/eller har en slap underlæbe. Hvis patienten drikker væske fra en ske, scores 1.

Score 2 hvis patienten ikke konsekvent tilpasser læbelukket til koppen/glasset. Måske ophører han med at tilpasse læbelukket til koppen/glasset undervejs i måltidet, eller han tilpasser kun lejlighedsvis læbelukket under måltidet.

Score 3 hvis patienten konsekvent tilpasser læbelukket til kop/glas under hele måltidet.

Kan drikke med almindeligt sugerør

Under dette item, observeres patientens evne til komplet læbelukke samtidig med, at der dannes et undertryk i munden. Hvis patienten ikke selv efterspørger et sugerør,

gives det til ham i løbet af undersøgelsen. Patienten behøver kun at tage nogle få sug af sugerøret og behøver ikke at bruge det under hele måltidet, med mindre han vælger det selv. Sugerøret bør have en standarddiameter på ca. ½ cm. Det vil lette undersøgelsen af patientens sugekraft, hvis sugerøret er klart eller hvidt og væsken er farvet. På den måde er det muligt at afgøre, hvor højt oppe i sugerøret væsken er, og med hvilken lethed eller besvær patienten kan suge væsken op.

Score 1 hvis patienten afviser at bruge sugerøret, ikke kan forestille sig, hvordan sugerøret bruges, putter sugerøret i munden, men uden at suge, ikke kan skabe nok sugekraft til at få væske op i sugerøret, eller kun kan få væsken delvist op i sugerøret. Hvis patienten ikke kan stoppe med at suge væske op, scores 1.

Score 2 hvis patienten med besvær lige netop får suget væske op til munden, eller hvis han har brug for hjælp og opmuntring for at udføre opgaven.

Score 3 hvis patienten jævnt og ubesværet suger væske op til munden.

Drikker af kop/glas uden der løber væske fra munden

Under dette item, observeres om der løber væske fra patientens mundvige eller fra underlæben, når han drikker af kop/glas.

Score 1 hvis der løber væske fra patientens mund næsten hver gang han drikker, eller hvis han indtager væske med en ske. Spild af store mængder væske vil det ofte resultere i, at trøjen, kraven eller halsudskæringen bliver våd, eller at der falder dråber på bordet.

Score 2 hvis der løber små til moderate mængder væske fra patientens mund, når han drikker, eller hvis det forværres i løbet af måltidet. Spild af små mængder væske vil som regel ikke resultere i, at tøjet bliver vådt, men der kan forekomme væske i mundvigene som kan løbe ned mod hagen. Hvis der ofte løber små mængder væske fra patientens mund, kan hagen til sidst blive våd. Spild af moderate mængder væske vil til sidst resultere i, at tøjet bliver vådt; men hvis dette kun sker en enkelt gang, vil det resultere i en våd hage.

Score 3 hvis der på intet tidspunkt løber væske fra patientens mund, når han drikker, eller hvis der kun løber få dråber væske fra patientens mund én eller to gange i løbet af måltidet.

Holder væsken i munden inden der synkes

Under dette item, observeres patientens evne til at forme en bolus af væsken og holde den væk fra læberne, samt evnen til at kontrollere og lukke læberne tæt nok til at undgå, at væsken løber ud af munden. Dette item scores, uanset om patienten drikker af kop, glas, ske eller med sugerør. Hvis der benyttes mere end én fremgangsmåde til at drikke, så observerer ergoterapeuten alle fremgangsmåderne. Patienten gives den laveste af disse scorer, men noter om én af fremgangsmåderne resulterede i, at der løb mindre væske fra patientens mund.

Score 1 hvis der løber store mængder væske fra patientens mund, når han har taget en tår. Spild af store mængder væske vil det ofte resultere i, at trøjen, kraven eller halsudskæringen bliver våd, eller at der falder dråber på bordet.

Score 2 hvis der løber små til moderate mængder væske fra patientens mund, når han har taget en tår. Spild af små mængder væske vil som regel ikke resultere i, at tøjet bliver vådt, men der kan forekomme væske i mundvigene som kan løbe ned mod hagen. Hvis der ofte løber små mængder væske fra patientens mund, kan hagen til sidst blive våd. Spild af moderate mængder væske vil til sidst resultere i, at tøjet bliver vådt; men hvis dette kun sker en enkelt gang, vil det resultere i en våd hage.

Score 3 hvis der på intet tidspunkt løber væske ud af patientens mund.

Kan drikke flere slurke ad gangen

Under dette item, observeres patientens evne til at koordinere drikkefunktionen og åndedrættet, så han kan tage mere end én mundfuld væske ad gangen, hver gang han løfter sit glas. At drikke med flere slurke ad gangen, defineres her som mere end én mundfuld væske, der indtages med et synk imellem uden at glasset flyttes fra læberne (glasset løftes, væske suges ind, væske synkes, væske suges ind, væske synkes, glasset sænkes). Dette observeres både, når der drikkes af kop eller glas og når der drikkes med sugerør. Patienten skal ikke have besked på at gøre dette. Hvis ikke patienten spontant drikker flere slurke ad gangen, kan det skyldes, at han kompenserer for et funktionstab.

Score 1 hvis patienten på intet tidspunkt drikker flere slurke ad gangen fra kop, glas eller sugerør, eller hvis patienten kun kan drikke flere slurke ad gangen én eller to gange i løbet af måltidet. Hvis patienten drikker væske med ske, scores 1.

Score 2 hvis patienten ikke konsekvent drikker flere slurke ad gangen fra kop, glas eller sugerør i løbet af måltidet, eller drikker flere slurke ad gangen med sugerøret og ikke fra koppen og glasset (eller omvendt).

Score 3 hvis patienten drikker flere slurke ad gangen med både kop, glas og sugerør.

Har uændret stemmekvalitet efter at have drukket

Under dette item, observeres patientens evne til at beskytte sine luftveje mod penetration og aspiration af væske. Ved penetration og aspiration, sker der en forandring af patientens stemmekvalitet. En hæs stemme er karakteriseret ved at blive dyb og få en grov kvalitet. En gurglende stemme er karakteriseret ved at lyde våd som om der er væske på stemmelæberne. For at afgøre om penetration og/eller aspiration forekommer i forbindelse med udtrætning, bør patienten observeres for dette under hele måltidet.

Score 1 hvis patienten mister sin stemme efter at have drukket, hvis han pga. fx afasi ikke kan udtrykke sig verbalt ved måltidets begyndelse, hvis han udviser ændret stemmekvalitet efter at have drukket små mængder væske eller hvis stemmekvaliteten ændres ved måltidets begyndelse. Ved et fuldstændigt tab af stemmen ved én enkelt lejlighed uden tilstedeværelse af anden abnormitet, scores 1.

Score 2 hvis patienten udviser ændret stemmekvalitet efter at have drukket en stor mængde væske, eller hvis stemmekvaliteten ændres mod slutningen af måltidet.

Score 3 hvis patienten på intet tidspunkt udviser ændret stemmekvalitet efter at have drukket.

Renser luftvejene, hvis der er behov efter indtagelse af væske

Under dette item, observeres patientens evne til at rense luftvejene efter at penetration eller aspiration af væske er sket. I denne funktionelle vurdering, er det uvæsentligt at skelne mellem en bevidst eller reflektorisk rensning af luftvejene. Derfor skal ergoterapeuten være opmærksom på samtlige af alle patientens forsøg på at hoste eller at rømme sig.

Score 1 hvis patienten ikke renser svælget effektivt; dvs., hvis patienten lyder som om han trækker vejret anstrengt eller kun genererer et svagt host eller rømmen. Hæsheden eller den gurglende stemme vil forsat være til stede efter forsøg på at rense svælget. Hvis åndedrættet eller talen høres med hæshed eller gurglende lyd og patienten ikke forsøger at rense svælget, scores 1.

Score 2 hvis patienten rømmer sig eller hoster under måltidet og det resulterer i en reduktion af stemmens hæshed eller gurglelyd.

Score 3 hvis patienten på intet tidspunkt behøver at rømme sig eller hoste under måltidet.

Skala for indtagelse af fast føde

Formålet med skalaen er at vurdere patientens evne til at indtage fast føde. Skalaen omhandler de forskellige motoriske færdigheder, der er forbundet med at spise samt de observerbare tegn på faryngeal dysfunktion og svækket luftvejsbeskyttelse. Patienten skal observeres i indtagelse af alle fødeemner med fast konsistens, inklusiv puré og budding.

Former og slutter overlæben tæt til bestik

Under dette item, observeres når patienten tager maden fra bestikket. Overlæben bør slutte tæt ned til bestikket for at fjerne maden. Overlæben skal slutte så meget til, at fx skeen renses, så der kun er en ubetydelig rest tilbage på den.

Score 1 hvis patienten ikke slutter overlæben tæt ned til skeen, hvis maden skrabes af med tænderne, eller hvis læben ikke konsekvent slutter fuldstændigt tæt til bestikket og det dermed resulterer i en anseelig restmængde på skeen eller gaflen. Hvis patienten mades og hjælperen ikke tillader, at patienten bruger overlæben, scores 1, men noter årsagen til denne lave score.

Score 2 hvis patienten bruger overlæben utilstrækkeligt; dvs., enten ikke bruger overlæben hver gang skeen eller gaflen skal tømmes, eller bruger overlæben med utilstrækkelig styrke til at rense fx skeen.

Score 3 hvis patienten konsekvent bruger overlæben tilstrækkeligt til at tømme og rense skeen eller gaflen.

Holder maden i munden inden der synkes

Under dette item, observeres patientens evne til at forme og kontrollere bolus i munden samt evnen til at kontrollere og lukke læberne tæt nok til at holde på maden. Der scores ud fra patientens udførelse gennem hele bearbejdelsen af bolus som forberedelse til selve synkningen.

Score 1 hvis patienten konsekvent taber mad fra munden, eller hvis der lejlighedsvis tabes store mængder af bolus.

Score 2 hvis patienten ofte eller lejlighedsvis taber små mængder mad fra munden. Det kan forekomme indimellem i løbet af måltidet, eller begynde at opstå mere konsekvent, når patienten udtrættes efterhånden som måltidet skrider frem.

Score 3 hvis patienten på intet tidspunkt taber mad fra munden i den orale fase.

Anvender et funktionelt tyggemønster

Under dette item, observeres patienten, når han indtager fødeemner, der skal tygges (fast konsistens, der er heterogen, hård, trævlet, blød, hakket). Observer, hvordan patienten bevæger kæben, når han tygger; dvs. tygger patienten med et roterende tyggemønster med både vertikale og cirkulære (malende) bevægelser af underkæben. Der scores i forhold til patientens udførelse ved samtlige observerede konsistenser, og patienten gives den laveste af disse scorer. Hvis bestemte konsistenser forårsager mere besvær end andre, noteres dette.

Score 1 hvis patienten ikke forsøger at tygge, eller hvis der kun observeres en suttebevægelse (tungen bevæges frem og tilbage). Hvis patienten kun ernæres med purékost, scores 1.

Score 2 hvis patienten kun bevæger kæben op og ned i en hakkende bevægelse, kaldet ”gumlen”. Det ser ud som om patienten tygger med et symmetrisk op/ned - bid på det samme punkt, igen og igen.

Score 3 hvis patienten konsekvent tygger med et roterende tyggemønster.

Tygger hensigtsmæssigt i forhold til fødeemner

Under dette item, observeres hvor mange gange patienten tygger hver mundfuld mad. Skønt de fleste tygger en mundfuld ca. 10 gange, så er der individuelle forskelle på, hvor mange gange, der er brug for at tygge før en mundfuld mad er findelt tilstrækkeligt. Dette er influeret af patientens kæbestyrke, tandstatus, bolusstørrelse samt fødeemnerne, der tygges. Hvis patienten tygger med åben mund, er det let at se om maden bliver tilstrækkeligt bearbejdet. Hvis patienten tygger med lukket mund, må du bruge din dømmekraft til at afgøre, hvorvidt fødeemnerne tygges tilstrækkeligt. Væske, budding og purékost bør ikke tygges.

Score 1 hvis patienten ikke forsøger at tygge, eller hvis han tygger en mundfuld uden at stoppe eller synke. Hvis patienten kun ernæres med purékost, scores 1.

Score 2 hvis patienten tygger enten overdrevent længe eller utilstrækkeligt, eller hvis patienten foruden at tygge faste fødeemner, også tygger væske, budding og/eller purékost.

Score 3 hvis patienten konsekvent tygger alle faste konsistenser hensigtsmæssigt.

Placerer bolus hensigtsmæssigt under tygning

Under dette item, observeres patientens evne til at lateralisere tungen og placere bolus så den kan tygges. Patienten bør placere bolus bagtil i munden på kindtænderne, uanset om han har tænder eller ej. Personer med få tænder vil placere bolus på de tænder i underkæben, der har de tilsvarende tænder i overkæben til at give modstand, når der tygges. Hvis disse tænder er fortænder, så vil tyggeeffektiviteten sandsynligvis være ringere end hos en person, der kan bruge sine kindtænder. Derfor scores der ud fra, hvor patienten anbringer bolus, uanset patientens tandsæt.

Score 1 hvis patienten placerer bolus mod den hårde gane, ikke former en sammenhængende bolus, eller ikke tygger. Hvis patienten kun ernæres med purékost, scores 1.

Score 2 hvis patienten indimellem placerer bolus på fortænderne eller kindtænderne.

Score 3 hvis patienten konsekvent placerer bolus på kindtænderne.

Mængde af madrester i munden efter synk

Under dette item, observeres patientens evne til at forme bolus til en sammenhængende enhed, synke og efterlade mundhulen ren uden madrester. Observer indersiden af patientens mund, når han åbner den for at tage den næste bid. Hvis det er vanskeligt at vurdere om der er madrester i munden, kan du bede om lov til at se ind i patientens mund.

Score 1 hvis mere end halvdelen af bolus resterer i munden. Hvis patienten ikke synker før han tager den næste bid, eller hvis patienten samler maden i kindhulerne og ikke synker det konsekvent, scores 1.

Score 2 hvis mindre end halvdelen af bolus resterer i munden, eller hvis der er antydning af madrester efter patienten har sunket. Dette kan være svært at se, hvis ikke ergoterapeuten sidder så hun kan se hele mundhulen. Madrester er ofte lokaliseret bagtil på tungen eller i kindhulerne.

Score 3 hvis munden er fuldstændig renset for madrester, efter patienten har sunket.

Madresters placering i munden efter synk

Under dette item, observeres madresternes placering i munden. Dette item kan observeres samtidig med det forrige item.

Score 1 hvis maden klæber fast til den hårde gane, eller hvis der er mad i kindhu-

lerne. Mad i kindhulerne observeres lettest ved at se på kindens kontur.

Score 2 hvis der er mad på eller omkring tungen eller på tænderne.

Score 3 hvis der ikke er madrester i munden efter patienten har sunket.

Synker uden anstrengelse

Under dette item, observeres hvor let patienten har ved at synke. Hvis patienten skal anstrenge sig for at synke, vil halsmusklerne formodentlig kontraheres og blive temmelig tydelige. Patienten kan have behov for at nikke med hovedet, når han synker, eller patientens synk kan høres meget tydeligt fra svælget.

Score 1 hvis patienten konsekvent skal anstrenge sig for at synke, når han spiser faste fødeemner af enhver slags.

Score 2 hvis patienten skal anstrenge sig for at synke, når han spiser visse faste fødeemner, eller indimellem behøver at anstrenge sig for at synke.

Score 3 hvis patienten på intet tidspunkt behøver at anstrenge sig for at synke.

Synker kun 1 eller 2 gange per mundfuld

Under dette item, observeres det antal gange patienten behøver at synke før munden er tilstrækkelig renset til den næste mundfuld. For at afgøre om antallet af synk er normalt, bør fødeemnets konsistens tages i betragtning. Et eller to synk anses for normalt. Mere end to synk tyder på besvær med bolustransport i munden.

Score 1 hvis patienten konsekvent behøver at synke mere end to gange for hver mundfuld.

Score 2 hvis patienten behøver at synke mere end to gange ved visse faste fødeemner, eller indimellem synker mere end to gange.

Score 3 hvis patienten på intet tidspunkt synker mere end to gange ved hver mundfuld.

Koordinerer åndedræt og spisning under måltidet

Under dette item, observeres patientens koordinering mellem spisning og åndedræt. Mange patienter har svært ved at trække vejret og tygge på samme tid, eller har svært ved at holde åndedrætspausen længe nok til at kunne synke. Disse problemer har en tendens til at forværres ved udtrætning. Observer patientens åndedrætsbevægelse flere gange under hele måltidet. Vær særlig opmærksom, når maden er inde i munden og når bolus synkes. Måske begynder patienten at trække vejret dybere, udånde kraftigere, stønne eller pruste, eller har en uregelmæssig åndedrætsrytme eller åndedrætsdybde.

Score 1 hvis patienten på et hvilket som helst tidspunkt under måltidet, får så svært ved at trække vejret, at læber, negle eller ansigt bliver blåligt (cyanotisk), hvis patienten har åndedrætsbesvær, der ikke forårsager cynaose, men som vedvarer under hele måltidet eller som indtræder omkring starten af måltidet.

Score 2 hvis patienten har åndedrætsbesvær uden tegn på cyanose, der indtræder mod slutningen af måltidet eller som optræder sporadisk i løbet af måltidet.

Score 3 hvis patienten opretholder en rolig og regelmæssig åndedrætsbevægelse under hele måltidet og ikke får åndenød på noget tidspunkt.

Har uændret stemmekvalitet efter at have spist

Under dette item, observeres patientens evne til at beskytte luftvejene mod penetration og aspiration af faste konsistenser. Ved penetration og aspiration sker der en forandring med patientens stemmekvalitet. En hæs stemme er karakteriseret ved at blive dyb og få en grov kvalitet. En gurglende stemme er karakteriseret ved at lyde våd, som om der er væske på stemmelæberne. For at afgøre om penetration og/eller aspiration forekommer i forbindelse med udtrætning, bør patienten observeres for dette under hele måltidet.

Score 1 hvis patienten mister sin stemme efter at have spist, hvis han fx pga. afasi ikke kan udtrykke sig verbalt ved måltidets begyndelse, hvis han udviser ændret stemmekvalitet efter at have spist små mængder faste konsistenser eller hvis dette sker tæt ved begyndelsen af måltidet. Ved et fuldstændigt tab af stemmen ved én enkelt lejlighed uden tilstedeværelse af anden abnormitet, scores 1.

Score 2 hvis patienten udviser ændret stemmekvalitet, efter at have spist en stor mængde fast konsistens, eller hvis stemmekvaliteten ændres mod slutningen af måltidet.

Score 3 hvis patienten på intet tidspunkt udviser ændret stemmekvalitet efter at have spist.

Renser luftvejene, hvis der er behov efter indtagelse af fast føde

Under dette item, observeres patientens evne til at rense luftvejene efter at penetration eller aspiration af faste fødeemner er sket. I denne funktionelle vurdering, er det uvæsentligt at skelne mellem en bevidst eller reflektorisk rensning af luftvejene. Derfor skal ergoterapeuten være opmærksom på samtlige af alle patientens forsøg på at hoste eller at rømme sig.

Score 1 hvis patienten ikke renser svælget effektivt; dvs. hvis patienten lyder som om han trækker vejret anstrengt eller kun genererer et svagt host eller rømmen. Hæsheden eller den gurglende stemme vil forsat være til stede efter forsøg på at rense svælget. Hvis åndedrættet eller talen høres med hæshed eller gurglende lyd og patienten ikke forsøger at rense svælget, scores 1,

Score 2 hvis patienten rømmer sig eller hoster under måltidet og det resulterer i en reduktion af stemmens hæshed eller gurglelyd.

Score 3 hvis patienten, på intet tidspunkt, behøver at rømme sig eller hoste under måltidet.

Skala for håndtering af konsistenser

Formålet med skalaen er at vurdere patientens evne til at håndtere forskellige faste og flydende konsistenstyper velvilligt og sikkert. Skalaen undersøger patientens dømmekraft, påpasselighed, følsomhed samt reaktion i forhold til forskellige konsistenstyper.

Hvis patienten ikke accepterer et bestemt fødeemne, vender han måske hovedet væk eller afviser at åbne munden. Nogle patienter opdager først, at de ikke er i stand til at håndtere en konsistenstype, når den er i munden. Hvis patienten spytter mad ud, virker ængstelig, græder eller på anden vis udviser tøven eller modstand overfor at synke, så bør dette tolkes som manglende evne til at håndtere denne konsistenstype. Nogle patienter tager måske én mundfuld og afviser derpå at spise mere. Dette bør tolkes som en ikke-accept, med mindre patienten giver mundtligt udtryk for, at han ikke kan lide smagen af maden.

Alle items i skalaen scores på samme måde ud fra nedenstående:

Score 1 hvis en konsistenstype ikke tilbydes under MISA-måltidet, hvis patienten på intet tidspunkt accepterer konsistenstypen, eller hvis patienten på intet tidspunkt kan håndtere konsistenstypen.

Score 2 hvis patienten indimellem accepterer konsistenstypen, eller hvis det vurderes at det indimellem er usikkert for patienten at indtage konsistenstypen.

Score 3 hvis patienten accepterer konsistenstypen, hver gang han får den præsenteret og ikke har noget besvær med at håndtere den.

Kategoriseringen af fødeemner bør ske i overensstemmelse med de følgende retningslinjer.

FASTE KONSISTENSER			
KATEGORI	**DEFINITION**	**EKSEMPLER**	
Heterogen/ Blandet	Mad med mere end 2 konsistenser, der er blandet. Det noteres om patienten har vanskeligheder ved en blanding af væske og fast konsistens, eller en blanding af 2 faste konsistenser.	• Kolde morgenmadsprodukter med mælk (cornflakes, havregryn, mysli) • Yoghurt med frugtstykker	• Suppe med grøntsager • Millionbøf med kartoffelmos • Boller i ris og karry • Rugbrød med skive pålæg/ost
Trævlet	Fortrinsvis kød, hvor kødfibrene måske falder nemt fra hinanden, men som skal tygges godt inden det synkes.	• Steg af oksekød, kalvekød, svinekød • Tilberedt fjerkræ	• Let kogte asparges • Frugt med skræl
Hård	Fødeemner, der ikke er trævlede, men kræver, at de bides over.	• Småkager / kiks • Knækbrød • Ristet brød • Tvebakker/skorper • Kammerjunker	• Rå eller let kogte grøntsager • Frugt uden skræl • Rugbrød / franskbrød uden kerner
Hakket/ granuleret	Fødeemner, som naturligt er i små stykker; eller fødeemner af anden konsistens, der er skåret i små stykker mindre end ½ cm.	• Ris • Bulgur • Solsikke-/ pinjekerner	• Hakket oksekød/millionbøf • Kødsovs • Hakket mad
Klæbrig	Fødeemner, der har en tendens til at klæbe til bestik eller indersiden af munden og svælget.	• Kartoffelmos • Leverpostej • Chokolade	• Flødeost / Brie / smelteost • Nutella
Blød	Fødeemner, der holder formen på tallerkenen og som kræver minimal tyggekraft.	• Pandekager • Pasta • Æg • Blød ost • Fisk	• Kogte grøntsager • Modne bananer • Henkogt frugt uden saft • Kage
Puré	Fødeemner fra de andre kategorier, der er pureret til en sammenhængende konsistens, der ikke skal tygges. Der kan være små klumper i. Puréen kan være let flydende, blød eller mere fast.	• Varm morgenmadsprodukt (Havregrød, ølle-brød, rismelsgrød o. lign) • Risengrød	• Pureret fuldkost • Frugtgrød / frugtmos
Budding	Fødeemner, der er puré-lignende i konsistens, men samtidig er fuldstændig jævn og blød.	• Budding • Mousse • Fromage • Kagecreme	• Græsk Yoghurt • Ymer / tykmælk • Grøntsags rand • Frugtgelé

VÆSKE KONSISTENSER			
KATEGORI	**DEFINITION**	**EKSEMPLER**	
Vand	Tyndtflydende	• Vand fra hane	• Flaskevand uden smag
Tynd væske	Væske med konsistens som vand, men med smag	• Bouillon • Klar suppe uden fyld • Frugtjuice (æble / appelsin) • Saftevand / læskedrik	• Mælk • Sorbet is • Italiensk is • Te/kaffe • Sodavand
Nektar	Væsker, der naturligt har en tykkere konsistens end vand + væsker, der er fortykket til denne konsistens	• Cremet legeret suppe • Tomat- /grøntsags juice • Ananas- /sveskejuice • Fløde is • Proteindrik	• Hyldebærsuppe • Kakaomælk • Kærnemælk /koldskål • Tyk milkshake • Æggesnaps • Fortykket væske
Honning	Vand + væsker, der er fortykket med fortykningsmiddel til en konsistens som flydende honning eller sirup, der løber meget langsomt.	• Fortykket vand • Fortykket væske • Sirup	• Cremesovs • Frugtsovs
Budding	Vand + væsker, der er fortykket med fortykningsmiddel til budding eller gelé konsistens (vær opmærksom på, at gelékost betragtes som tynd væske, fordi det smelter i munden).	• Gelékost • Fortykket vand	• Fortykket tynd juice

Oversigten er adapteret fra Pardoe (1993)[2].

[2] O.A. Oversigten er adapteret til dansk madkultur af diætister på Herlev Hospital.

Appendix C - Item location and fit statistic for the six MISA-DK subscales

The appendix C presents the individual item fit after the extended Rasch analysis on each individual MISA-DK subscale (Table IC. Item location and fit statistic for the six MISA-DK subscales).

1. **The positioning subscale:** all items were initially consistent with Rasch model expectations and all were retained.
2. **The Self-feeding skills subscale:** item 9 manifested non-uniform DIF by gender and was removed from the scale.
3. **The liquid ingestion subscale:** local item dependency was present for items 12/14/15 and 17/18, and they were combined into two testlets.
4. **The Solid ingestion subscale**: local item dependency was present for items 21/23, 24/25, 26/27, and 29/30, and they were combined into four testlets.
5. **The Texture management-solids subscale**: items were not consistent with Rasch model expectations and all were regarded as single items.
6. **The Texture management liquid subscale**: items were not consistent with Rasch model expectations and all were regarded as single items.

Table IC. Item location and fit statistic of the six MISA-DK subscales

Subscales	Loc	SE	FR	χ^2	df	*P*	*F*	df1,df2	P
Positioning scale									
1. Maintain symmetry of posture	-0.74	0.21	-0.17	2.01	2	0.365	0.73	2,89	0.487
2. Maintain adequate head position for feeding	-0.82	0.22	-1.08	3.92	2	0.141	2.91	2,89	0.060
3. Maintain 90-degree hip angle	-1.46	0.23	0.36	3.43	2	0.180	1.77	2,89	0.178
4. Maintains postural stability in the trunk	3.03	0.24	-0.74	0.41	2	0.814	0.46	2,89	0.630
Self-feeding skills scale									
5. Able to grasp utensil/food-item functionally and bring it to the mouth	-0.20	0.21	-1.15	4.76	2	0.092	3.22	2,90	0.045
6. Able to grasp cup/glass functionally and bring it to the mouth	-0.50	0.21	-1.36	1.71	2	0.425	0.56	2,90	0.571
7. Selects appropriate utensil for food item	-0.62	0.21	-1.10	1.71	2	0.412	0.86	2,90	0.428
8. Takes appropriately-sized mouthfuls	0.44	0.21	-1.40	6.25	2	0.044	5.05	2,90	0.009
9. Able to focus on meal	Single item								
10. Demonstrates good judgment	0.23	0.20	1.79	0.92	2	0.633	0.49	2,90	0.614
11. Able to complete the meal without fatigue	0.65	0.22	1.37	5.36	2	0.068	2.50	2,90	0.088
Liquid ingestion scale									
Testlet of: 12. Seals lips on cup/glass + 14. Prevents leakage of liquid from cup/glass while drinking +15.Prevents leakage of liquid from mouth before swallow	-1.40	0.11	-0.55	1.36	2	0.509	0.26	2,88	0.775
13. Able to draw liquid from a standard straw	-0.59	0.18	-0.62	4.23	2	0.121	2.99	2,88	0.056
16. Able to take a sequence of sips	1.51	0.18	2.00	0.70	2	0.703	0.28	2,88	0.759
Testlet of: 17. Demonstrates same voice quality after drinking + 18.Clear the airway if necessary after liquids	0.48	0.11	0.22	0.86	2	0.654	0.52	2,88	0.596
Solid ingestion scale									
19. Close upper lip on utensil	-1.25	0.22	-0.27	0.74	2	0.690	0.51	2,97	0.604
20. Prevents the loss of food from the mouth before swallowing	-1.59	0.23	1.32	0.33	2	0.850	0.13	2,97	0.878
Testlet of: 21. Use functional chewing pattern + 23. Positions bolus when chewing	0.33	0.12	0.58	1.29	2	0.524	0.55	2,97	0.576
22. Chewing appropriate to food item	0.93	0.18	-0.25	4.52	2	0.104	2.79	2,97	0.066
Testlet of: 24. Quantity of food remaining in mouth after swallow + 25. Location of food remaining in the mouth after swallow	0.44	0.12	2.22	5.68	2	0.059	2.55	2,97	0.084
Testlet of: 26. Swallow without extra effort + 27. Swallows only once or twice per mouthful	-0.51	0.14	-1.11	3.49	2	0.175	3.12	2,97	0.049
28. Maintains respiratory pattern throughout meal	0.87	0.16	1.58	3.50	2	0.174	1.27	2,97	0.286
Testlet of: 29. Demonstrate same voice quality after eating +30. Clear the airway if necessary after solids	0.78	0.11	1.51	7.52	2	0.023	3.52	2,97	0.033
Texture management - solids									
Capable of eating heterogeneous, fibrous, hard, minced/granular, sticky, soft, puree and pudding solids	8 single items								
Texture management - liquids									
Capable of drinking water, thin juices, nectar-, honey-, and pudding consistency liquids	5 single items								

Abbreviations: Loc, location expressed in logits; SE, Standard error; FR, Fit residual; $\chi 2$,chi-square; df, degrees of freedom; F, F-statistics.

Paper I

Scandinavian Journal of Occupational Therapy. 2011; 18: 282–293

informa
healthcare

ORIGINAL ARTICLE

Content validation of a Danish version of "The McGill Ingestive Skills Assessment" for dysphagia management

TINA HANSEN[1], HEATHER C. LAMBERT[2] & JENS FABER[3]

[1]*Occupational Therapy, Herlev University Hospital, Herlev, Denmark,* [2]*School of Physical and Occupational Therapy, McGill University, Montreal, Quebec, Canada, and* [3]*Department of Endocrinology, Herlev University Hospital, Herlev, Denmark*

Abstract

This study addresses the first steps in the cross-cultural adaptation of a Danish version of the McGill Ingestive Skills Assessment (MISA), which quantifies eating and drinking abilities by scoring a meal observation. The original Canadian MISA was translated and adapted into Danish (MISA-DK). For content validation of the MISA-DK, a judgemental quantification process was applied using 13 experts. Thereafter, the MISA-DK was pilot tested by 16 occupational therapists. Finally, the MISA-DK was linked to the International Classification of Functioning, Disability and Health (ICF). Content validity of 43 items was found for 93% in terms of adequacy, 67% in terms of clarity of item description, 86% in terms of clarity of score descriptions, and 93% in terms of relevance. Thirteen of 14 sections of the instruction manual and score sheet were content valid. In light of these results, a revised MISA-DK was produced for the pilot test, which then found content validity for all sections and 98% of the items. The ICF linking resulted in 41 ICF-categories, which may reflect the complexity of eating and drinking as well as a multidimensional structure of the MISA-DK. In conclusion, the MISA-DK is prepared for psychometric testing using classical as well as modern test theory.

Key words: *cross-cultural validation, eating and swallowing disorders, ICF, observation-based assessment, occupational therapy*

Introduction

Eating and drinking are complex basic activities of daily life, which require effective, coordinated function of the motor, sensory- and cognitive system (1–3). These activities are strongly influenced by the context (cultural, social, physical, personal, spiritual, and temporal) surrounding a meal routine and are essential to health and well-being (4). However, age-related physiologic changes in the aero-digestive tract in conjunction with various medical conditions may cause eating and drinking problems in older people leading to dysphagia, i.e. eating and swallowing disorders (2,5–7). Dysphagia in the elderly is associated with increased comorbidity and mortality (2,5–7) as well as reduced quality of life (7–9).

The goal of occupational therapy within dysphagia management is to enable safe and independent eating, drinking, and swallowing (1,4). This necessitates specific assessment of all the phases of the eating, drinking, and swallowing process (1,4). These inter-dependent phases are conceptualized as: the pre-oral phase where food/liquid is brought to the mouth; the oral phase where food/liquid is prepared and formed into a bolus for transportation into the pharynx; the pharyngeal phase were the bolus is transported to the oesophagus simultaneous with airway protection to prevent aspiration; and the oesophageal phase where the bolus moves from the oesophagus to the stomach (1–4). However, no evidence-based and dysphagia-specific clinical assessments based on the conceptual foundations of occupational therapy are

Correspondence: Tina Hansen, MSc OT, Department of Occupational Therapy, Herlev University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark. E-mail: ergotp.tina@gmail.com

(Received 5 April 2010; revised 23 June 2010; accepted 2 September 2010)

ISSN 1103-8128 print/ISSN 1651-2014 online
DOI: 10.3109/11038128.2010.521949

currently available in Denmark (10). If the contributions of occupational therapy to health care are to be explicit, the focus must be on occupational performance (11,12). In order to define and describe the occupational performance of the patient in the eating, drinking, and swallowing process, occupational therapy assessment may involve observation during a mealtime in the patient's habitual context (1,4). Assessment based on observation necessitates a clinical assessment instrument that has been developed through a rigorous methodology with established evidence of validity and reliability (1,11,12). Additionally, the instrument should investigate all areas of possible influences on a problem and should be able to inform decisions about appropriate interventions (12). Recently, the Canadian Association of Occupational Therapists has published "The McGill Ingestive Skills Assessment" (MISA), developed by Lambert et al. (13). The MISA is designed to capture aspects of eating and drinking not included in technical assessments of dysphagia or traditional swallowing trials (13). These procedures usually require the administration of only a few spoonfuls of a limited variety of liquid or food textures in a standardized and artificial environment (2,14–16). What is interpreted from these assessments does not necessarily predict actual performance in a natural meal activity, and may lead to interventions with limited relevance for the patient (2,14). The MISA seems to provide an alternative approach as it evaluates the ability of elderly patients to consume a variety of foods and liquids safely and independently during the usual mealtime routine, and guides the occupational therapist in identifying areas where skills are impaired and amenable to rehabilitation (13). The items in the MISA have been generated from an extensive literature review and focus-group methodology (17). Pilot testing and preliminary psychometric testing were carried out to enable item reduction and refinement (17). Finally, large-scale testing of the MISA's psychometric properties indicates adequate construct validity, known-groups validity, predictive validity, internal consistency, interrater reliability, and intrarater reliability (18,19). Thus, the MISA may be of value for Danish occupational therapists practising dysphagia management. However, as the original Canadian version of the MISA is in English, a translation and cross-cultural adaptation process is necessary to assure the assessment's content validity, as well as its uniform administration and interpretation across different languages and cultures (20,21).

Content validation is a critical step in the translation process (20). Several aspects of test quality are defined within the concept of content validity, namely construct definition, adequacy, clarity, and relevance (22). This implies that every element of a clinical assessment instrument is to be evaluated (23). A traditional procedure in content validation involves subject-matter experts whose judgements are quantified via formalized scaling procedures (22–24). After a judgemental evaluation and possible modifications, a pilot test is then used to identify potential problems related to the clarity of the instructions and the wording of the items when used by those who would administer the instrument (20). Finally, by linking a new clinical assessment instrument to the International Classification of Function, Disability and Health (ICF) (25), the relationship between the assessment items and the theoretical definition of the construct they aim to measure can be examined (26,27). Thus, the content validity can be established further (20). A precise understanding of the content of a new clinical assessment instrument may guide researchers and practitioners when choosing an assessment instrument as well as facilitate the direction for further validation in the development process (26).

The objective of this study was to translate and culturally adapt the MISA into a Danish version (MISA-DK), as well as to content validate the translated version. The specific aims in the content validation process were to investigate whether any items or sections of the MISA-DK needed modification in order to be adequate, clear, and relevant; to investigate whether the instructions and items in the revised MISA-DK appeared clear when used in clinical practice; and to investigate to what extent the content of the revised MISA-DK represents the ICF.

Material and methods

The study was carried out in four phases from December 2008 to February 2010:

(1) translation and adaptation of the MISA;
(2) judgemental evaluation of the content validity of the MISA-DK;
(3) pilot testing of the revised MISA-DK;
(4) linking the content of the revised MISA-DK to the ICF.

Instrument

The MISA consists of a four-page score sheet and an instruction manual, which outlines the conceptual framework, the specific procedures for administering and scoring, and the evidence of MISA's reliability and validity. The MISA is administered during the observation of a test meal with 13 different food and liquid consistencies. However, individual food preferences or dietary restrictions are taken into account.

The MISA is composed of 43 items distributed in five subscales (Figure 1):

(1) a positioning scale assessing the patient's ability to maintain a position that is safe for eating and drinking;
(2) a self-feeding skills scale assessing the patient's self-feeding skills, behaviour, and judgement;
(3) a liquid ingestion scale assessing the patient's oral motor skills for liquids;
(4) a solid ingestion scale assessing the patient's oral motor skills for solids;
(5) a texture management scale assessing the patient's ability to manage a variety of food textures.

Each item is described in detail in the instruction manual, and is scored on a three-point ordinal scale. Scores of 1 and 3 represent the absence or the presence of the specific functional performance, and a score of 2 represents deficient or inconsistent functional performance. For the exact scoring of each item within the first four scales, the categories in the three-point ordinal scale are further described in the instruction manual. For the texture-management scale, the items represent a categorization of various solid and liquid consistencies and the scoring is based on the patient's ability to manage the different consistencies willingly and safely. The scores are summed to give subscale scores and a total score for the entire assessment (13).

Translation and adaptation

The translation method was based on Geisinger (20) and Douglas & Craig (28). Initial forward translation of the MISA was independently carried out by three translators (two certified translators with no knowledge of dysphagia, and one bilingual occupational therapist experienced within the field of dysphagia and a native speaker of Danish). A synthesis of the three translations was performed by a review committee of two occupational therapists, a dietician, and the first author (TH) (all experienced within the field of dysphagia, bilingual, and native speakers of Danish). The review committee scrutinized and compared all translations with the original English version of the MISA. Care was taken to focus on the conceptual rather than the literal equivalence, and emphasis was on semantic equivalence across languages, conceptual equivalence across cultures, and translational quality (20,21,28). In this process, some of the examples of the categorized consistencies in the texture-management scale were adapted into Danish food culture. In order to ensure semantic equivalence and that no essential information had been lost, the consensus version was compared with the original version of the MISA by a bilingual occupational therapist who is a native speaker of English (USA). This resulted in some minor changes to the wording of several items and score descriptions. During the whole translation process, the primary author of the original version of the MISA (Heather C. Lambert) was consulted when needed. Finally, the MISA-DK was proofread by a teacher in Danish.

Participants

Judgemental evaluation of the content validity. The MISA-DK was judged by 13 experts recruited purposively from five main hospitals in the Capital Region of Denmark. All experts were certified occupational therapists and experienced within the field of dysphagia for at least one year. The average length of time since graduation in occupational therapy was 7.6 years (range 2–28); the average length of clinical experience within the field of dysphagia was 6.1 years (range 2–17); and 54% had participated in postgraduate education in dysphagia.

Pilot testing. The MISA-DK was judged by 16 pilot testers, eight of whom also participated in the judgemental evaluation of the content validity. The pilot testers were recruited purposively from seven main hospitals and three rehabilitation centres in the Zealand region of Denmark. All pilot testers were certified occupational therapists, practised in dysphagia management, and had the opportunity to use the MISA-DK in their own clinical setting. The average length of time since graduation in occupational therapy was 6.5 years (range 2–28); the average length of clinical experience within the field of dysphagia was 5.0 years (range 1–17); and 50% had participated in postgraduate education in dysphagia.

Linking to the ICF. The MISA-DK was linked to the ICF independently by two occupational therapists (ICF raters) recruited from the Danish ICF network (29), a multidisciplinary society which communicates knowledge within the field of the ICF. Both ICF raters were experienced in using the ICF as a tool in their professional work.

All occupational therapists and the participating patients in the pilot study gave informed consent. The study was approved by the local ethical committee in the Capital region (Reg. No: H-C-2009-061) as well as the Danish Data Protection Authority (Reg. No: 2009-41-3719).

Positioning scale	Liquid ingestion scale	Texture management scale
1. Maintain symmetry of posture	12. Seals lips on cup/glass	31. Capable of eating heterogeneous textures
2. Maintain adequate head positioning for feeding	13. Able to draw liquid from a standard straw	32. Capable of eating fibrous solids
3. Maintain 90-degree hip angle	14. Prevents leakage of liquid from cup/glass while drinking	33. Capable of eating hard solids
4. Able to sit upright without leaning on arm	15. Prevents leakage of liquid from mouth before swallowing	34. Capable of eating minced/granular solids
	16. Able to take a sequence of sips	35. Capable of eating sticky solids
Self-feeding skills scale	17. Demonstrates same voice quality after drinking	36. Capable of eating soft soldis
5. Able to grasp utensil functionally and bring it to the mouth	18. Demonstrates clear airway after liquids	37. Capable of eating puree
6. Able to grasp cup/glass functionally and bring it to the mouth	**Solid ingestion scale**	38. Capable of eating pudding
7. Selects appropriate utensil for food item	19. Close upper lip on utensil	39. Capable of drinking water
8. Takes appropriately-sized mouthfuls	20. Prevents the loss of food from the mouth before swallowing	40. Capable of drinking thin juices
9. Able to focus on meal	21. Use functional chewing pattern	41. Capable of drinking nectar consistency liquids
10. Demonstrates good judgment	22. Chewing appropriate to food item	42. Capable of drinking honey consistency liquids
11. Tolerates physical effort of meal	23. Position bolus when chewing	43. Capable of drinking pudding consistency liquids
	24. Quantity of food remaining in mouth after swallow	
	25. Location of food remaining in the mouth after swallow	
	26. Swallow without extra effort	
	27. Swallows only once or twice per mouthful	
	28. Maintain respiratory pattern throughout meal	
	29. Demonstrates same voice quality after eating	
	30. Demonstrates clear airway after solids	

Note: Each item is scored on a 3-point ordinal scale. Score 1 = the absence of the specific functional performance. Score 2 = deficient or inconsistent functional performance. Score 3 = the presence of the specific functional performance (13).

Figure 1. Scales and items included in the McGill Ingestive Skills Assessment (MISA) (13).

Procedure

Judgemental evaluation of the content validity. As recommended (24,30), the experts were provided with the conceptual basis for the MISA via a two-hour introduction meeting and the MISA-DK was handed over. The experts were asked to examine the MISA-DK and to respond independently to a validity questionnaire (24,30) within three weeks. The validity questionnaire was divided into two parts. Part one covered the adequacy of the item terms in reflecting the item content, the clarity of the item and score descriptions, and the relevance of each item. Part two covered the clarity of the sections in the instruction manual and the sections of the score sheet. For each content validity domain, a four-point Likert scale was used (24): 1 = not at all adequate/clear/relevant, 2 = needs major modifications to be adequate/clear/relevant, 3 = needs minor modifications to be adequate/clear/relevant, 4 = very adequate/clear/relevant. The experts were given the opportunity to provide open-ended comments. The results of the judgement were presented and discussed with the experts at a two-hour follow-up meeting. All suggestions on modifications were sent to the primary author of the original version of the MISA for final approval.

Pilot testing. The pilot testers attended a one-day training programme in the use of the MISA-DK. Subsequently, they applied the revised MISA-DK to at least five patients at their own facility, and answered the clarity domain of the validity questionnaire concerning the MISA-DK.

Linking to the ICF. The ICF raters were introduced to the revised MISA-DK and the linking rules (26). Each ICF rater independently identified and extracted all meaningful concepts within the overall purpose of the MISA-DK and all 43 items, inclusive of the item and score descriptions. Each meaningful concept was then linked to the most precise ICF category within the ICF components: *body functions* (b), *body structures* (s), *activities and participation* (d), *environmental factors* (e), and *personal factors* (26,27). The ICF categories are represented by the letters b, s, d, and e, and are followed by a numerical code at different levels. An example selected from the *body function*s (b) component is given in Figure 2.

If a single item encompassed different concepts, the information in each concept was linked separately. For example, the two meaningful concepts *symmetry of posture* and *reposition after weight shift* were identified for the item "maintain symmetry of posture", and were linked to the ICF categories *d4153 maintaining a sitting position* and *d4106 shifting the body's centre of gravity*. Concepts that could not be linked to the ICF because of insufficient information were labelled "*nd*" (not definable). If a concept was not contained in the ICF classification, then this concept was labelled "*nc*" (not covered by the ICF) (26,27). The ICF raters were asked not to use the "*other specified*" and "*other unspecified*" ICF categories, and were asked to document additional information if concepts were difficult to link, and if they were not definable or not covered by the ICF.

Analyses

Judgemental evaluation of the content validity. In order to estimate quantitative evidence of content validity of the MISA-DK, the Content Validity Index (CVI) (24,31) and the Average Deviation (AD) Index (32) were used.

The CVI indicates the proportion of experts who gave ratings of 3 or 4 on the content validity questionnaire, i.e. endorsed an item or section as adequate/clear/relevant. CVI values can range from 0 to 1 (24). For this study, a universal agreement approach was applied (33). This implied that items or sections of the MISA-DK achieving CVI = 1 were deemed to be content valid; otherwise they needed to be scrutinized for possible modifications. As the CVI is associated with a risk of chance agreement among the experts and there is a loss of information collapsing the four-point Likert scale responses into two nominal categories (33), a second analysis of interrater agreement was undertaken using the AD index (32). The AD index is proposed as a measure of interrater agreement for ratings on a Likert scale of a single target on a single occasion (32). The AD index is calculated by determining the extent to which each expert's rating differs from the mean or the median rating, summing up the absolute values of these deviations and dividing by the number of deviations (32). As the four-point Likert scale is based on an ordinal scale construction (34) the median rating (AD_{Md}) was applied in this study. Burke & Dunlap (32) set the upper cut-off limit for acceptable and statistical significant agreement levels as a function of the sample size and the number of categories on the Likert scale ($\alpha = 0.05$). Accordingly, the upper cut-off level was 0.65, indicating acceptable AD_{Md} results unlikely to be obtained by chance (32).

For all the items and the sections of the MISA-DK, the CVI values were examined to determine whether they were endorsed by the experts or not; then the AD_{Md} values were examined to determine the level and significance of agreement among the experts. Thereafter, the open-ended comments from the experts were considered to determine possible modifications of the MISA-DK.

Code	Category	Level
b5	Functions of the digestive, metabolic and endocrine systems	first
b510	Ingestion functions	second
b5105	Swallowing	third
b51051	Pharyngeal swallowing	fourth
b51058	Swallowing, otherspecified	fourth
b51059	Swallowing, unspecified	fourth

Figure 2. Example of ICF codes and categories at different levels from the body functions component (25).

Pilot testing. The pilot testers' judgements of the clarity domain of the MISA-DK were also evaluated using the CVI (24,31) and the AD_{Md} (32). For the CVI, the universal agreement approach was applied (33), and for the AD_{Md}, an upper cut-off level of 0.67 was used (32). Differences in the judgements between the pilot testers who participated in the judgemental evaluation versus those who did not was analysed using the Mann–Whitney U-test and a two-sided significance level of 0.05 (34).

Linking to the ICF. Consensus between the two ICF raters was used to decide which ICF categories should be linked to the MISA-DK (27). If there was disagreement between the selected categories in terms of the specific level, the less specific higher-level category was selected, as this level incorporates the attributes from the more specific lower-level categories (25). In case of absolute disagreement between the two ICF raters, TH made a decision based on the additional information documented by the ICF raters. If the same ICF category was addressed repeatedly in a single item the category was counted only once. The consent density was analysed using the average number of identified concepts per item and the content diversity was examined using the number of ICF categories per concept (35). For the content density, a value exceeding 1 indicates that more than one concept was identified. For the content diversity, a value of 1 indicates that each concept was linked to a different ICF category and a value below 1 indicates that several concepts were linked to one and the same ICF category (35). In addition, the frequency of the linked ICF categories that were attributed to the ICF components was calculated.

All statistical calculations were carried out using SAS 9.1 and SPSS 17.0.

Results

Judgemental evaluation of the content validity

The results of the judgemental evaluation of the content validity of the MISA-DK are presented in Table I. Of the 43 items on the MISA-DK, adequate content validity (i.e. CVI = 1.00) was found for 40 items in terms of adequacy of the item term, for 29 items in terms of clarity of the item description, for 37 items in terms of clarity of the score descriptions, and for 40 items in terms of relevance. For all 43 items, $AD_{Md} < 0.65$ indicated acceptable and statistically significant agreement among the experts in terms of adequacy of the item terms and clarity of the item and the score descriptions. When considering the relevance of the items by means of the AD_{Md}, there were acceptable and statistically significant agreement levels for all but one item, which obtained an AD_{Md} value of 0.69 indicating that this result could have been obtained by chance. In total, the content validity domains not endorsed by means of the CVI referred to 21 items of which 13 items belong to the texture-management scale. The comments made by the experts for these items were that it seems that they contain several purposes. Other comments made by the experts were specific suggestions for alteration of item terms as well as linguistic modifications of items and score descriptions.

Of the seven sections in the instruction manual of the MISA-DK, adequate content validity (i.e. CVI = 1.00) was found for all sections in terms of clarity. $AD_{Md} < 0.65$ indicated that acceptable and statistical significant agreement was obtained in all cases. Of the seven sections on the score sheet, the CVI values indicated that the experts did not endorse one section, "the summing up section", in terms of clarity with a CVI = 0.92. It was pointed out that the

Table I. Content validity of the MISA-DK judged by experts (n = 13).

	CVI[a] (range)	CVI = 1.00 (number of items/sections)	AD_{Md}[b] (range)	$AD_{Md} < 0.65$ (number of items/section)
Items (n = 43)				
Adequate item terms	0.92–1.00	40	0.00–0.54	43
Clear item descriptions	0.92–1.00	29	0.00–0.31	43
Clear score definitions	0.85–1.00	37	0.00–0.53	43
Relevant items	0.85–1.00	40	0.00–0.69	42
Instruction manual (n = 7)[c]				
Clear sections	1.00–1.00	7	0.00–0.46	7
Score sheet (n = 7)[d]				
Clear sections	0.92–1.00	6	0.08–0.54	7

Notes: [a]CVI = Content Validity Index: CVI of 1.00 reflects endorsement by all 13 experts. [b]AD_{Md} = Average Deviation Index based on median ratings: $AD_{Md} < 0.65$ are acceptable and statistically significant ($\alpha = 0.05$). [c]Seven sections: conceptual framework, using the MISA, intended use, preparation, test meal, set-up, and scoring. [d]Seven sections: summing up section, positioning scale, self-feeding skills scale, liquid ingestion scale, solid ingestion scale, and the texture management scales for solids and liquids.

equations for the calculation of the percentage score were difficult to interpret. $AD_{Md} < 0.65$ indicated that acceptable and statistically significant agreement was obtained in all cases.

Based on the results and with the approval of the Canadian MISA's primary author, certain modifications were applied to the MIDA-DK (see Table II). No items or sections were eliminated.

Pilot testing

The results of the pilot test of the revised MISA-DK are presented in Table III. Adequate content validity (i.e. CVI = 1.00) was found for 42 of the 43 items in terms of clarity of item and score descriptions. AD_{Md} results <0.67 indicated acceptable and statistically significant agreement levels among the pilot testers in terms of clarity of all item and score descriptions. In total, the content validity domains not endorsed by the CVI referred to item 3, "maintain 90-degree hip angle". No specific comments were made by the pilot testers for this item. Adequate content validity (i.e. CVI = 1.00) was found for all sections in the instruction manual and the score sheet of the MISA-DK in terms of clarity. AD_{Md} results < 0.67 indicated that acceptable and statistically significant agreement was obtained in all cases.

There was no significant difference in the ratings of the items or sections of the revised MISA-DK between the pilot testers who participated in the judgemental evaluation versus those who did not (the calculated U-values ranged from U = 19.5 to U = 32.0, and the *p*-values ranged from $p = 0.063$ to $p = 1.0$, Mann–Whitney U-test ($\alpha = 0.05$)).

Table II. Approved modifications of the MISA-DK.

Altered item terms	Item 4 "Maintains postural stability in the trunk"
	item 11 "Able to complete the meal without fatigue"
	Item 18 "Clear the airway if necessary after liquids"
	Item 30 "Clear the airway if necessary after solids"
Modification of item descriptions	Item 30, and items 31–43
Modification of score descriptions	Item 3, item 13, item 15, item 17, item 18, and item 30
Modification of sections	In the section "Scoring" in the instruction manual, the equations for the percentage score were elaborated.

Linking to the ICF

A total of 41 different ICF categories were addressed in the MISA-DK of which 60% could be selected on the basis of absolute consensus between the two ICF raters. The overall purpose of the MISA-DK was linked to the ICF categories *d550 eating* and *d560 drinking* within the activity and participation component.

The results of the ICF linking process at item level are presented in Table IV. For the 43 items of the MISA-DK a total of 214 concepts were identified; 117 of these concepts were identified for the 13 items in the texture-management scale. These items are described and scored in a similar manner, which resulted in identical concepts across items. In general, the density ratio of 5 indicates that several concepts

Table III. Content validity of the MISA-DK judged by pilot testers (n = 16).

	CVI[a] (range)	CVI = 1.00 (number of items/sections)	AD_{Md}[b] (range)	$AD_{Md} < 0.67$ (number of items/sections)
Items (n = 43)				
Clear item descriptions	0.94–1.00	42	0.00–0.31	43
Clear score definitions	0.81–1.00	42	0.00–0.56	43
Instruction manual (n = 7)[c]				
Clear section	1.00–1.00	7	0.00–0.25	7
Score sheet (n = 7)[d]				
Clear section	1.00–1.00	7	0.13–0.20	7

Notes: [a]CVI = The Content Validity Index: CVI of 1.00 reflects endorsement by all 16 pilot testers. [b]AD_{Md} = Average Deviation Index based on median ratings: $AD_{Md} < 0.67$ are statistically significant (α =.05). [c]Seven sections: conceptual framework, using the MISA, intended use, preparation, test meal, set-up, and scoring. [d]Seven sections: summing up section, positioning scale, self-feeding skills scale, liquid ingestion scale, solid ingestion scale, and the texture-management scales for solids and liquids.

were identified for all 43 items. The diversity ratio of 0.2 indicates that concepts from different items were linked to the same ICF category. For example, the ICF category *d440 fine hand use* was linked to concepts within two items in the self-feeding skills scale, and the ICF category *b5103 manipulation of food in the mouth* was linked to concepts within four items on the solid ingestion scale. Six concepts could not be linked to the ICF and were coded *nd* or *nc*. For the 41 linked ICF categories, 63.4% were within the body functions component, 2.4% were within the body structures component, 24.4% were within the activity and participation component, and 9.8% were within environmental factors. No items were linked to personal factors.

The most frequently addressed categories were related to the ingestion functions (*b510*) at second, third, and fourth level in chapter *b5* "*functions of the digestive, metabolic and endocrine system*". In total, concepts from 28 items were linked to categories in this chapter. The other linked categories were related to *mental functions (b1)*, *sensory functions and pain (b2)*, *voice and speech functions (b3)*, *respiration functions (b4)*, *neuromusculoskeletal and movement-related functions (b7)*, *structures involved in voice and speech (s1)*, *learning and applying knowledge (d1)*, *mobility (d4)*, *self-care (d5)*, *interpersonal interactions and relationships (d7)*, *products and technology (e1)*, and *support and relationship (e3)*.

Table IV. Frequencies of items, concepts, and ICF categories in the MISA-DK.

Number of items (n)	43
Number of concepts (n)	214
Content density (concepts per item)	5
Number of different ICF categories	41
Content diversity (categories per concept)	0.2
Concepts not covered or defined by the ICF (n)	6
ICF categories per component	
Body functions (n (%))	26 (63.4)
Body structures (n (%))	1 (2.4)
Activity and participation (n (%))	10 (24.4)
Environmental factors (n (%))	4 (9.8)

Discussion

Discussions of results

The objective of this study was to translate and culturally adapt the MISA into Danish and to examine the content validity of the translated version using judgemental evaluation, pilot testing, and linking to the ICF. In the translation phase, only a small number of adaptations were made because of cultural motives, and it may be assumed that the underlying concept of the Canadian MISA is appropriate for use in Denmark. For further cross-cultural adaptation it will, however, be suitable to compare how the items function across Canadian and Danish groups, which can be realized through statistical methods such as differential item function analysis (20,21) or structural equation models (36).

As no former content validity study using quantitative methods has been performed on the MISA, the results from this study cannot be compared with other research. In the present content-validation process, the extent to which the items and the sections in the MISA-DK were adequate, clear, and relevant using expert judgements was quantified by means of the CVI and the AD index. For the CVI, a universal agreement approach (33) was applied, which resulted in modification and adaptation of 21 scale items and one section of the MISA-DK. However, it may be questioned whether using a universal agreement

approach is too rigorous as the likelihood of achieving total agreement is decreased when the number of experts is high (31). According to Lynn (24), CVI values equal to 0.78 are appropriate when the number of experts exceeds 10. Using this guideline would have resulted in no modification or adaptation of the MISA-DK. The additional interrater agreement analyses using the AD index indicated acceptable and statistically significant agreement levels for all content-validity domains but one in terms of relevance for one item. This may signal that the performed modifications and adaptations based on the CVI results may have been superfluous. However, the open-ended comments made by the experts and the discussions at the follow-up meeting did reflect a need for modification and adaptation. This highlights the necessity of a rigorous standard when interpreting the CVI (33) as well as the importance of combining a quantitative and a qualitative approach in the content-validation process (22,23). Furthermore, the pilot test of the revised MISA-DK came up with 98% of the items obtaining CVI = 1.00 for the clarity of the item descriptions and the score definitions versus 67% and 86% in the judgemental evaluation of the content validity. This may well reflect that the MISA-DK did improve from the provided modifications and adaptations. Although the relevance of three items was not endorsed by the experts using the CVI, no items were eliminated. It is suggested that for a scale as a whole to be judged as having excellent content validity it should be composed of items of which 80% obtain CVI values that meet the stated criteria (30,31), which was CVI = 1.00 in this study. The judgemental evaluation of the content validity in this study may indicate that the items in the MISA-DK form a strong scale in terms of relevance, as 93% of the items obtained CVI = 1.00. In addition, before any deletions of items in the MISA-DK it is necessary to investigate the internal scale validity. This may be achieved using modern test theory models such as Rasch analyses (20,21,37).

In the pilot test, item 3 did not meet the stated CVI criterion in relation to the clarity of the item and score description. As the AD index indicated acceptable and statistical significant agreement among the pilot testers for item 3, and as no specific comments were made, no modifications were carried out. As well, the MISA-DK is now composed of items of which 98% obtained CVI = 1.00 for the clarity domain, which indicates excellent content validity for this domain (30,31). However, it is important to investigate further how this item functions. This may also be achieved using modern test theory models such as Rasch analyses (20,21,37).

In the content-validation process, the extent to which the content in the MISA-DK is represented by the different ICF components was examined. Most of the identified meaningful concepts in the MISA-DK were covered by the ICF model. Not surprisingly, the purpose of the MISA-DK was linked to the ICF categories *d550 eating* and *d560 drinking* from the activity and participation component. Going into details at item level, a more varied picture of the content in the MISA-DK was reflected, which may indicate that the MISA-DK covers most of the ICF components. This is in line with Treats (14), who argues that all ICF components should be emphasized in dysphagia management in order to reduce the risk of non-compliance with the dysphagia interventions. However, as the present linking process resulted in a greater representation of the body functions component compared with the activity and participation component, it may be assumed that the MISA-DK assesses underlying functions for occupational performance (11,12). It is important to consider that a complete view of occupational performance must cover performance skills and performance patterns in conjunction with patient factors (i.e. body functions and body structures), activity demands, and contexts (1,11,12). In this view, the MISA-DK should not be used in isolation, but must be supplemented by assessment instruments using the patient's perspective (12). Within the area of dysphagia, specific assessments using the patient's perspective have become available (38).

The MISA-DK was linked to categories across four ICF components. This may indicate that the MISA-DK integrates the complexity of eating and drinking. However, it may also indicate that the MISA-DK is multidimensional (39). It is important to realize that when the scores of items measuring different components are added to form one overall score and these scores are based on an ordinal scale construction, interpretation of the final result and the real meaning of the finding may be questionable (37,39). If we want to estimate quantities from the counts of observed behaviours, one of the theoretical requirements for an assessment instrument is that it is based on a unidimensional construct (39). An approach for investigating the dimensional structure and scalability can be Rasch analyses (37,39), which will need to be applied to the MISA-DK in future research. The development of the original version of the MISA is based on the classical test theory of reliability and validity. As validation of an assessment instrument is an ongoing process (37), application of Rasch analyses to the original version of the MISA would also be beneficial. In addition, as the MISA was developed before the publication of the ICF (17) it would be appropriate to repeat the linking process on the Canadian version of the MISA. This would contribute to further

cross-cultural adaptation of the MISA-DK as well as verifying the results from this study.

Although the four ICF components are represented in the MISA-DK, the diversity ratio was relatively low, and the ICF categories related to ingestion functions in the body function component were frequently addressed. The primary evaluations of these functions are the fibreoptic endoscopic evaluation of swallowing (FEES) or the videofluoroscopic modified barium swallows evaluation (VFS) (2,14–16). Therefore, the FEES or the VFS should be considered as criterion standards in further validation of the MISA-DK. Likewise, as the MISA-DK addresses a broad number of functions including mental and mobility-related functions, tools addressing these issues will also have to be applied.

It seems that the results of the ICF linking process may deepen the understanding of the experts' judgements of the MISA-DK. Several items were linked to the same ICF categories, which may explain why some items were judged by the experts not to be relevant. Several meaningful concepts were identified for the items in the texture-management scale, which may explain the experts' perception of several purposes for these items. However, this did not seem to be an issue for the pilot testers, who participated in a one-day training programme in the use of the MISA-DK. This may underline the necessity of formal training in the use of the MISA-DK. For the moment this is not a prerequisite, but it is recommended that the occupational therapist becomes acquainted with the items and their scoring before administration of the assessment (13).

In general, the results of this content-validation process emphasize the importance of a comprehensive approach when evaluating the content of a clinical assessment. The procedures used to facilitate and evaluate content validity are to be based on both judgemental methods and statistical evaluations (22). Using statistical evaluation based on modern test theory models such as Rasch analyses will provide further information regarding content and construct validity (22) as well as cross-cultural validity (20,21). This will be applied in the forthcoming field testing of the MISA-DK in addition to testing its reliability and validity using classical test theory.

Discussions of methodology

The translation procedure applied in this study used an alternative approach to the standard back-translation methodology. Back-translations have no clear scientific basis to prevent the production of poor translations (20,28). Alternative approaches have been successful in producing adaptations that are of equal psychometric quality to the original versions of health-related quality-of-life questionnaires (40). After the forthcoming psychometric testing of reliability and validity of the MISA-DK, comparison with the results from the large-scale testing of the Canadian MISA (19) will verify whether or not the applied translation and content-validation method was successful (20,21).

For the analysis of the content validity, the CVI and the AD index were used. Other agreement indices such as Kappa statistics (41) have been suggested (33). However, very large samples of experts are required to obtain a high degree of agreement with a high degree of confidence when using Kappa (33,41). In addition, the value of Kappa depends on the prevalence in each category used for the ratings (41), and it can therefore be misleading despite high proportions of agreement on adequate content validity (31).

In repeated judgements of the content validity of an instrument, the experts should be equally qualified (24,30). This was not the case in this study as the pilot testers attended a formal training course in the MISA-DK and the experts in the judgemental evaluation did not. In that sense, the pilot testers may have gained a deeper understanding of the content in the MISA-DK. Furthermore, some of the pilot testers also participated in the judgemental evaluation of the MISA-DK, which may have influenced their second judgements. However, no statistically significant differences in the judgements were found between the pilot testers who participated in the judgemental evaluation versus those who did not.

A potential limitation for the linkage procedure is that, although being very familiar with the ICF classification, the ICF raters did not know about the linking rules a priori. It is important to note that the linking rules, although being standardized guidelines, present some challenges in establishing meaningful concepts and consensus in the linking process (27). Therefore, before the linking procedure, a practice period with the linking rules would have been preferable. Furthermore, the reliability of the linking process would have been strengthened by higher number of ICF raters.

In the present study, the content-validation process involved occupational therapists at all levels. This might have weakened the range of representation and the expertise in the content-validation process. Dysphagia is typically managed by a multidisciplinary team of speech-language pathologists, dieticians, physicians, radiologists, nurses, and respiratory therapists (1,2). However, in Denmark occupational therapists are primarily responsible for the management of dysphagia (10). Therefore, it was decided to base the content validation solely on experts from the field of

occupational therapy. In addition, using purposeful sampling for the expert panel and the pilot testing ensured a range of clinical expertise and postgraduate education in dysphagia.

Conclusion and implications for further research and practice

This preliminary study represents the first steps in the cross-cultural adaptation of the Danish version of the MISA. Experts' judgement of content validity suggested that the direct use of the MISA-DK was inappropriate and that some modifications and adaptations were needed. Thereafter, pilot testing provided strong evidence for the content validity of the revised MISA-DK. Furthermore, by linking it to the ICF it was found, that the MISA-DK covers most of the ICF components and reflects the complexity of eating and drinking activities. The results of the ICF linking process also reflected a need to apply modern test theory models such as Rasch analyses in addition to classical test theory in the validation process. It seems now that the MISA-DK is ready for further field testing in order to verify its psychometric properties in terms of reliability and validity using classic test theory (20) as well as internal scale validity, dimensionality, scalability. and differential item function using modern test theory (20,21,39).

Acknowledgements

The authors are grateful to all who participated in this study. The study was funded by the Department of Occupational Therapy at Herlev University Hospital.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Meriano C, Latelle D. Occupational therapy interventions: Function and occupation. Thorofare, NJ: Slack; 2008.
2. Cichero JYA, Murdoch BE. Dysphagia: Foundation, theory and practice. Chichester: Wiley; 2006.
3. Leopold NA, Kagel MA. Dysphagia – ingestion or deglutition? A proposed paradigm. Dysphagia 1997;12:202–6.
4. AOTA. Specialized knowledge and skills in feeding, eating, and swallowing for occupational therapy practise. Am J Occup Ther 2007;61:686–99.
5. Schindler S, James MD, Kelly MD. Swallowing disorders in the elderly: State of the art review. Laryngoscope 2002;122: 589–602.
6. Achem SR, DeVeault KR. Dysphagia in aging. J Clin Gastroentero 2005;39:357–71.
7. Miller N, Carding P. Dysphagia: Implications for older people. Rev Clin Gerontol 2007;17:177–90.
8. Chen PH, Golup JS, Hapner ER, Johns MM. Prevalence of perceived dysphagia and quality-of-life impairment in a geriatric population. Dysphagia 2009;24:1–6.
9. Eslick GD, Talley HJ. Dysphagia: Epidemiology, risk factor and impact on quality of life – a population-based study. Aliment Pharmacol Ther 2008;27:971–9.
10. Kjærsgaard, A. Undersøgelse af synkeproblemer hos senhjerneskadede patienter (Assessment of dysphagia in patients with acquired brain damage) 2002. Available at: http://www.vfhj.dk/admin/write/files/824.pdf.
11. Dunn W. Measurement issues and practices. In: Law M, Baum C, Dunn W, editors. Measuring occupational performance: Supporting best practice in occupational therapy. 2nd ed. Thorofare, NJ: Slack; 2005. p. 21–32.
12. Fisher AG. Occupational Therapy Intervention Process Model: A model for planning and implementing top-down, client-centred, and occupation-based interventions. Fort Collins, CO: Three Star Press; 2006.
13. Lambert HC, Gisel EG, Wood-Dauphine S, Groher ME, Abrahamowicz M. McGill Ingestive Skills Assessment: User's manual. Ottawa: Canadian Association of Occupational Therapists; 2006.
14. Treats TT. Use of the ICF in dysphagia management. Seminars in speech and Language 2007;28:323–33.
15. Matino R, Pron G, Diamant N. Screening for oropharyngeal dysphagia in stroke: Insufficient evidence for guidelines. Dysphagia 2000;15:19–30.
16. Swigert NB. Update on current assessment practices for dysphagia. Top Geriatr Rehabil 2007;23:185–96.
17. Lambert HC, Gisel EG, Groher ME, Wood-Dauphine S. McGill Ingestive Skills Assessment (MISA): Development and first field test of an evaluation of functional ingestive skills of elderly persons. Dysphagia 2003;18: 101–13.
18. Lambert HC, Abrahamowicz M, Groher ME, Wood-Dauphine SM, Gisel EG. The McGill Ingestive Skills Assessment predicts time to death in an elderly population with neurogenic dysphagia: Preliminary evidence. Dysphagia 2005; 20:123–32.
19. Lambert HC, Gisel EG, Groher ME, Abrahamowicz M, Wood-Dauphine S. Psychometric testing of the McGill Ingestive Skills Assessment. Am J Occup Ther 2006;60:409–19.
20. Geisinger KF. Cross-cultural normative assessment: Translation and adaptation issues influencing normative interpretation of assessment instruments. Psychol Assess 1994;4: 304–12.
21. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: The universalist approach. Qual Life Res 1998;7:323–35.
22. Sireci SG. The construct of content validity. Social Indicators Res 1998;45:83–117.
23. Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: A functional approach to concepts and methods. Psychol Assess 1995;3:238–47.
24. Lynn MR. Determination and quantification of content validity. Nurs Res 1986;35:82–5.
25. WHO. ICF: International classification of functioning, disability and health. Geneva: World Health Organization; 2001.
26. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: An update based on lessons learned. J Rehabil Med 2005;37:212–18.
27. Xiong T, Hartley S. Challenges in linking health-status outcome measures and clinical assessment tools to the ICF. Arch Physio 2008;10:152–6.

28. Douglas SP, Craig CS. Collaborative and iterative translation: An alternative approach to back translation. J Int Mark 2007; 15:30–43.
29. Danish ICF network via Hanne Melchiorsen, Marselisborg-Centret (Homepage on the internet) (cited 21 September 2009). Available at: http://www.marselisborgcentret.dk/kontakt-os/forskning-udvikling/hanne-melchiorsen/#c1569.
30. Grant JS, Davis LL. Focus on quantitative methods: Selection and use of content experts for instrument development. Res Nurs Health 1997;20:269–74.
31. Polit DE, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health 2007;30:459–67.
32. Burke MJ, Dunlap WP. Estimating interrater agreement with the Average Deviation Index: A user's guide. Organ Res Methods 2002;5:159–72.
33. Beckstead JW. Content validity is naught. Int J Nurs Stud 2009;46:1274–83.
34. Portney IG, Watkins MP. Foundations of clinical research: Application to practice. 3rd ed. Upper Saddle River, NJ: Prentice Hall Health; 2009.
35. Stamm T, Gey S, Cieza A, Machold K, Kollerits B, Kloppenburg M, et al. Measuring function in patients with hand osteoarthritis – content comparison of questionnaires based on the International Classification of Function, Disability and health (ICF). Rheumatology 2006;45: 1134–1541.
36. Beckstead JW, Yang CY, Lengacher CA. Assessing cross-cultural validity of scales: A methodological review and illustrative example. Int J Nurs Stud 2008;45:110–19.
37. Dekker J, Dallmaijer AJ, Lankhost GJ. Clinimetrics in rehabilitation medicine: Current issues in developing and applying measurement instruments 1. J Rehabil Med 2005;37: 193–201.
38. McHorney CA, Robbins J, Lomax K, Rosenbek J C, Chignell K, Kramer AE, et al. The SWAL-QOL and SWAL-CARE outcomes tool for oropharyngeal dysphagia in adults, III: Documentation of reliability and validity. Dysphagia 2002;17:97–114.
39. Tesio L. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. J Rehabil Med 2003; 35:105–15.
40. Hedin PJ, Mckenny SP, Meads DK. The Rheumatoid Arthritis Quality of life (RAQol) for Sweden: Adaptation and validation. Scand J Rheumatol 2008;35:117–23.
41. Sim J, Wright CC. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. Phys Ther 2005;85:257–68.

Paper II

Disability & Rehabilitation, 2012; 34(10): 859–868
Copyright © 2012 Informa UK, Ltd.
ISSN 0963-8288 print/ISSN 1464-5165 online
DOI: 10.3109/09638288.2011.624249

informa
healthcare

RESEARCH PAPER

Validation of the Danish version of the McGill Ingestive Skills Assessment using classical test theory and the Rasch model

Tina Hansen[1], Heather C. Lambert[2] & Jens Faber[3]

[1]Department of Occupational Therapy, Herlev University Hospital, Herlev, Denmark, [2]School of Physical and Occupational Therapy, McGill University, Montreal, Canada, and [3]Department of Medicine/Endocrinology, Herlev Hospital, Herlev, Denmark

***Purpose*: The study aimed to validate the Danish version of the Canadian the "McGill Ingestive Skills Assessment" (MISA-DK) for measuring dysphagia in frail elders. *Method*: One-hundred and ten consecutive older medical patients were recruited to the study. Reliability was assessed by internal consistency (Chronbach's alpha). External construct validity (convergent and known-groups validity) was evaluated against theoretical constructs assessing the complex concept of ingestive skills. Internal construct validity was tested using Rasch analysis. *Results*: High internal consistency reliability with Chronbach's alpha of 0.77–0.95 was evident. External construct validity was supported by expected high correlations with most of the constructs related to ingestive skills ($r_s = 0.53$ to $r_s = 0.66$). The MISA-DK discriminated significantly between known-groups. Fit to the Rasch model (x^2 (*df*) = 12 (12), p = 0.424) and unidimensionality of the MISA-DK was confirmed after resolving disordered thresholds for 11 items and adjustment of local dependency. *Conclusion*: The psychometric properties of the MISA-DK equal the original Canadian version. Assessment of internal construct validity indicated multidimensionality due to local dependency. Although achieving good fit to the Rasch model after adjustments, additional studies are needed to establish cross-cultural validity. Finally, establishment of the inter- and intra-rater reliability of the MISA-DK is also needed.**

Keywords: Construct validity, Denmark, dysphagia, frailty, McGill Ingestive Skills Assessment, occupational therapy

Implications for Rehabilitation

- Validity evidence is a prerequisite to verify whether a measurement instrument in fact accomplish what it is supposed to accomplish.
- Using classical test theory in combination with the Rasch Model provides comprehensive insight of validity evidence.
- The Danish version of the McGill Ingestive Skills Assessment provides valid estimates of dysphagia patients' ingestive skill abilities.

Introduction

Dysphagia is a predictor of pneumonia in frail elders [1,2] and is associated with poor rehabilitation outcomes and reduced quality of life [3]. Frailty is characterised by vulnerability, general susceptibility to disease and poor outcome [4]. As dysphagia can produce impaired swallow efficacy and safety with subsequent malnutrition and aspiration pneumonia [1–3,5], it requires an attentive multidisciplinary dysphagia management approach in frail elders [5,6].

The goal of occupational therapy in dysphagia management is to assist patients to return to efficient and safe performance in eating and drinking activities [7]. Occupational therapists consider performance in activity as a dynamic interaction of the activity, the person and the environment [8], and observation of this interaction is a core part of the occupational therapy assessment process [8,9]. This necessitates clinical measurements with established evidence of validity and reliability [9,10].

In a recently published literature review concerning evidence-based clinical measures of elderly dysphagic patients' performance in eating and drinking during a natural meal [10], the McGill Ingestive Skills Assessment (MISA) [11] was recommended.

The MISA measures the ability of elderly patients to eat and drink safely and independently during the usual mealtime routine [11]. The conceptualisation of eating and drinking in the MISA is based on a construct termed "Ingestion" [11]. Ingestion includes cognition, physiological factors such as hunger, exteroceptive sensation of the meal, neck and truncal position, the manual and oromandibular aspects of

Correspondence: Mrs. Tina Hansen, Occupational therapist, Department of Occupational Therapy 53P1, Herlev University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark. Tel: +45 44 88 38 18. E-mail: ergotp.tina@gmail.com

(Accepted September 2011)

eating and drinking as well as the voluntary, automatic and reflex components of bolus preparation and the swallow [12]. The MISA assigns ordinal scores to the ingestive skills of the patient and provides subscale scores and a total score [11]. The items in the MISA have been generated and psychometrically tested using classical test theory, and shows adequate construct validity, predictive validity, internal consistency, inter- and intra-rater reliability [10,13–15].

The MISA has been translated into Danish [16] and establishment of its measurement equivalence (i.e. validity and reliability) is now required [17]. During the translation and adaptation of the Danish version of the MISA (MISA-DK) [16], it was suggested to analyse the MISA-DK using the Rasch model [18] as well. The Rasch model is useful for testing whether items from a scale measure a unidimensional construct, which is required for summation of ordinal scores [18,19]. Thus, the aims of this study were to investigate the measurement equivalency of the MISA-DK using classical test theory and investigate whether the MISA-DK appears to represent an unidimensional construct. The specific objectives were as follows:

1. Examination of the reliability and the external construct validity of the MISA-DK in terms of internal consistency, convergent and known-groups validity.
2. Examination of the internal construct validity of the MISA-DK in terms of fit to the Rasch model.

Methods

Design

We applied a cross-sectional design combining classical test theory [20] with the Rasch model [21]. This combination enables a comprehensive assessment of the capacity of the instrument to measure the intended construct, i.e. validity [20,22].

Classical test theory examines validity via theoretical assumptions about the targeted construct and its relations to other variables [20]. As no gold standard for measuring ingestive skills exists [10], it was not possible to use criterion validation. Therefore, we examined the relationship between the MISA-DK and constructs with known relationship to ingestive skills; i.e. construct validity [20]. Construct validity was examined by convergent validity and known-group validity [20]. Because ingestion is complex [12], the best constructs for convergent validation was determined to be: cognition measured with the Mini-Mental Status Examination (MMSE) [23], physical function measured with the Barthel-100 index (BI) [24], orofacial function measured with the Nordic Orofacial Test-Screening (NOT-S) [25] and swallowing function measured with the Water Swallow Test (WT) [26]. For the known-group validity, we examined whether the MISA-DK would be able to discriminate among groups with different levels of disability [27] in terms of frailty and pneumonia.

The Rasch model provides a mathematical model, which is a probabilistic form of Guttman scaling [18,21]. For the MISA-DK to be valid, it is expected that each item has its own level of difficulty on the trait (e.g. ingestive skills) and every patient has his or her own level of ability on the trait [21,28]. A more able patient will more likely succeed items than a lesser able patient, and easier items are more likely to be passed by all patients [21]. Thus, it is expected that the items forming a scale is unidimensional [18–22,29]. Unidimensionality includes that the items in the scale represent one common underlying latent construct, which is the only factor that accounts for variations in score patterns [18–22]. The presence of differential item function (DIF) and/or local dependency violate the requirement of unidimensionality [18]. DIF occurs when the probability of being rated on a particular score is not conditioned on the trait but is dependent of external factors such as gender or age [18,20,21]. Local dependency occurs when the score on an item depends on the score on other items after controlling for the latent trait or because of multidimensionality [29].

Participants

Patients consecutively admitted to two departments of general medicine at a large hospital in the Capital Region of Copenhagen between December 2009 and February 2011 were screened for inclusion within 48 hours of admission. The patients were invited to participate in the study if they were over 65 years, were not terminally ill, would require more than 2 days of hospitalization and were able to give personal information and written informed consent. The patients were excluded if they did not fulfil five criteria for direct swallowing evaluation [30], namely the ability to: remain alert for at least 15 minutes, sit in a chair or bed in at least a 60° upright position, swallow saliva, cough voluntary and clear the throat twice. Of 439 eligible patients, 168 patients were unable to give personal information and written informed consent and 87 patients declined. Of the remaining 184 patients, 74 (40%) were unable to perform the five swallowing criteria. This resulted in the inclusion of 110 patients. The study was approved by the local ethical committee in the Capital region (Reg. No: H-C-2009-061) and the Danish Data Protection Authority (Reg. No: 2009-41-3719).

Measurements

The MISA-DK is composed of 43 items distributed into six subscales: 1) positioning (4 items) addressing the patient's ability to maintain a position that is safe for eating and drinking; 2) self-feeding skills (7 items) addressing the patient's self-feeding skills, behaviour and judgement; 3) liquid ingestion (7 items) addressing the patient's oral motor and pharyngeal skills for liquids; 4) solid ingestion (12 items) addressing the patient's oral motor and pharyngeal skills for solids; 5) texture management-solids (8 items) addressing the patient's ability to manage eight solid food textures and 6) texture management-liquids (5 items) addressing the patient's ability to manage five liquid textures. Each item is scored on a three-point ordinal scale (1 = absent, 2 = inconsistent and 3 = present functional performance). High scores indicate high ability levels in ingestive skills [11,16].

The MMSE measures seven domains of cognition (temporal orientation, spatial orientation, immediate memory, attention

and calculation, recall, language, and visual construction). The score range from 0 to 30, and increasing scores indicate higher cognitive ability [23,31].

The BI measures the patient's performance in 10 activities of daily life. The items are related to self-care (feeding, grooming, bathing, dressing, bowel and bladder care and toilet use) and to mobility (ambulation, transfers and stair climbing). The score range from 0 to 100, and increasing scores indicate higher physical function [24]. The BI was routinely completed by the facility nurse staff or by interview to the patient.

The NOT-S is a screening instrument of orofacial dysfunction and contains a clinical examination with six domains (the face at rest, nose breathing, facial expression, masticatory muscle and jaw function, oral motor function and speech). The score range from 0 to 6, and higher score indicates orofacial dysfunction [25].

The WT included two stages. In stage 1, a teaspoon (5 ml) of water was given three times and those patients safe on at least two of three attempts were given a larger volume (60 ml) of water to drink continuously from a cup. The criteria for safety completion of stage 1 and 2 were: no delay or absence of up and forward laryngeal movement on attempted swallow, no cough or choking during or after the swallow, no change in voice quality and no signs of respiratory distress. Failure at either stage was recorded as a failed WT [26].

Frailty: The patients were considered frail if they fulfilled three or more of the following criteria [32]: unintentional weight loss, exhaustion, weakness, slowness and poor physical activity. The presence of *weight loss* was determined by the initial screening of the Nutrition Risk Screening [33] routinely performed and documented by the facilities' nursing staff. *Exhaustion* was measured through interview with the Danish version of the WHO-five Well-Being index (WHO-5). The score range from 0 to 100 and a cut off <50 indicate poor well-being [34]. W*eakness* was measured by decreased grip strength using a handheld dynamometer (average of 3 measures using dominant hand) and established norms at age and gender [35]. *Slowness* was measured as a time of >19 seconds on the "Timed Up & Go" test [36]. *Poor physical activity* was determined by a BI score <50, indicating moderate to severe functional disability [24].

Pneumonia: The presence of pneumonia was determined on basis of the diagnosis made by the medical physician of the patient and documented in the patients' medical file. Clinical findings, laboratory data, chest x-ray and antibiotic treatment were registered.

Procedure

The first author (TH), who is a senior occupational therapist with specialised knowledge and skills in dysphagia management, administered the MISA-DK to the patients at breakfast or lunch time. All additional measurements and data collection was performed within 2 days after the MISA-DK by a research assistant (RA), who is an experienced occupational therapist. TH and RA were blinded to the results of the additional measurements and the MISA-DK respectively. Before enrolment of patients, RA practised all the additional measurements with 10 patients under supervision of TH.

Statistical analysis

Reliability and external construct validity: The internal consistency reliability was analysed using Cronbach's alpha (α) [27]. The external construct validity was analysed using nonparametric statistics as the Kolmogorov-Smirnov test displayed not-normal distributions for several of the variables. For the convergent validity, Spearman's rho (r_s) was used [27]. We expected the MISA-DK total scale to correlate strongly with the MMSE, the BI, the NOT-S and the WT. For the subscales, we expected that: positioning correlated with the BI; self-feeding skills correlated with the BI and the MMSE; and solid and liquid ingestion as well as texture management correlated with the NOT-S and the WT. The magnitude for a strong correlation was set to >0.50 [27]. In order to assess the relative importance and the contributions of the convergent variables to variance in the construct (i.e. ingestive skills), stepwise multiple regression analysis was applied [20]. Analysis was conducted separately for the MISA-DK total scale and the subscales as dependent variables and the convergent variables as independent variables. Evidence of multicolinearity was not present and age should not be controlled for. For the known-group validity, the Mann Whitney U-Test and a two-sided significance level of 0.05 [27] were used to test whether the MISA-DK scales would discriminate between frail patients versus not frail patients, and between patients with and without pneumonia. The statistical analysis was undertaken using SPSS, version 17.0 (SPSS Inc., Chicago, IL).

Internal construct validity: The Rasch model specifies that the probability of a patient succeeding an item is a logistic function of the difference between the patients ability level and the difficulty of the item [18,21]. Thus, the ordinal scores are transformed into logits (log-odd units) [21]. Item and patient parameter are estimated separately and are placed on the same logit-scale centered by a mean item location of zero. Positive values reflect difficult items and high ability levels, and negative values reflect easy items and low ability levels [21]. Rasch analysis is an iterative process and a number of tests are performed [18,37–39], which we applied into four steps. All items of the MISA-DK were treated as one scale.

In step 1, three overall model fit statistics were considered. Two are item-person interaction statistics, which are a summary of all the individual item and person fit residuals (i.e. the degree of divergence between the Rasch model expectations and that actually accounted for in the raw data set). The fit residuals are transformed to approximate a *z*-score and represent a standardised normal distribution. For model fit, these summary fit residuals should have a mean close to 0.0 and a standard deviation (SD) of 1.0 [38], though SD <1.4 is usually accepted. The third fit statistic is an item-trait interaction statistic calculated as a chi-square (χ^2), which should be nonsignificant ($p > 0.05$). This fit statistic reflects whether the hierarchical ordering of the items is consistent across different levels of the trait (i.e. class intervals) [38]. The reliability of the scale using the person-separation index (PSI) was also considered. The PSI is analogue to Cronbach's α, except that it is calculated from the logit scale person estimates [38]. A PSI of 0.7 is a minimum acceptable level [40]. For further

examination of model fit, individual item and person fit-statistics by means of fit residuals, χ^2 and F-statistics were used. Individual item and person fit residuals between ±2.5 or χ^2 and F-statistic probability values above the Bonferroni adjusted α value of 0.05 were considered adequate model fit [37]. Large positive fit residuals indicate multidimensionality and large negative fit residuals indicates local dependency and redundancy [38].

In addition to the model fit statistics, the ordering of the score categories by means of the thresholds was investigated using the thresholds estimates and category probability curves [38]. Thresholds refers to the point between two adjacent score categories where either score is equally probable, and for a good fitting model, monotonicity is expected [21,37,38]. DIF was checked with regard to gender (male, female) and age groups (defined by the median of 83 years). DIF is detected via analysis of variance for each item [38]. Local dependency was investigated by inspecting the residual correlation matrix of the items [38]. Local dependency was evident by item residual correlations >0.2 above the average of all item residual correlations. Unidimensionality was examined using t-tests to compare person estimates derived from the two most disparate subsets of scale items [39]. The subsets were created based on principal component analysis of the residuals, and items with the highest positive and negative loadings on the first residual factor were used to construct the two subsets [39,41]. For a scale to be considered unidimensional, no more than 5% of cases should show a significant difference between their scores on the two subsets. If this is the case, a binomial test is used to calculate a 95% CI around the t-test estimate. Unidimensionality is supported if the value of 5% falls within the 95% CI [18].

After step 1, a fitting solution with continuous check of the above-mentioned points was sought. In step 2, disordered thresholds were resolved by combining adjacent categories, which may improve overall model fit [18,21,38]. In step 3, examination of misfitting items or DIF items was carried out, and a stepwise removal was considered. This solution is stopped at the point when good overall fit is achieved or no individual items displays misfit [21]. In step 4, local dependency was emphasised. This was dealt with by grouping the involved items into a testlet (a higher-order item) [18,42]. In this way, it is examined whether the local dependency is cancelled out at the test level. Finally, the targeting of the study sample was confirmed [18,37]. The Rasch analysis was performed using RUMM2030 [38]. As the MISA-DK is based on an ordinal scale construction, the polytomous version was applied [21]. A likelihood ratio test ($\chi^2 = 323$, $p < 0.001$) revealed that the partial credit model should be used [28,38]. As the score 1 for nine items was not represented by >1% [43], an adjustment for null categories was employed [38].

Sample size

For multiple regressions, the number of independent variables should not exceed the square root of the sample size [44]. This is 10,5 in this study, which is well above the number of included independent variables. For the Rasch analysis, a reasonable targeted sample of 100 patients will provide 95% confidence that the estimated item difficulty is within ±0.5 logits [45].

Results

Participants

The sample of 110 patients was represented by 50% males and females, respectively. The mean age was 81.9 (SD 7.6) years. The patients had on average 2.15 admission diagnoses (SD 1.1) and on average 2.7 chronic medical conditions (SD 1.6). The main diagnostic characteristics were distributed as follows: 63% had diseases of the circulatory system and 25% had sequellae after stroke, 57% had diseases of the respiratory system (chronic obstructive lunge disease and/or asthma) and 44% had a diagnosis of pneumonia, 35% had diseases of the musculoskeletal system, 25% had diabetes mellitus, 16% had urinary tract infection and 10% had diseases of the nervous system such as Parkinson's disease or epilepsy. The results of the MISA-DK scales and the validation variables are presented in Table I.

Reliability and external construct validity

Cronbach's α ranged from 0.77 to 0.95 for the MISA-DK scales (Table II, left column). On the whole, the correlations between the MIDA-DK scales and the convergent variables were

Table I. Distribution of MISA-DK scores and the scores of the validation variables.

	Mean (SD)	Sample range
MISA-DK (n = 110)		
Positioning scale	9.4 (2.1)	5–12
Self-feeding skills scale	17.2 (3.3)	8–21
Liquid ingestion scale	17.5 (3.1)	9–21
Solid ingestion scale	28.3 (5.5)	12–36
Texture management solids	17.9 (4.3)	8–24
Texture management liquids	12.7 (2.7)	6–15
MISA-DK total scale	102.9 (17.1)	58–128
Validation variables		
BI (n = 110)	48.8 (31.4)	0–100
MMSE (n = 102)	22.0 (5.4)	6–30
NOT-S (n = 102)	2.8 (1.5)	0–6
WT (n = 105)	Frequency (%)	
WT stage 1 failed	30 (28%)	
WT stage 2 failed	45 (43%)	
WT succeeded	30 (28%)	
Frailty criteria		
Unintentional weight loss (n = 105)	38 (36%)	
WHO-5 < 50 points (n = 105)	70 (67%)	
Weakness (n = 100)	63 (63%)	
TUG >19 seconds (n = 104)	78 (75%)	
BI <50 points (n = 110)	54 (49%)	
Frailty index (n = 104)		
Not frail	40 (38%)	
Frail	64 (62%)	
Pneumonia (n = 110)	48 (44%)	

MISA-DK, the Danish version of the McGill Ingestive Skills Assessment [11,16]; MMSE, the Mini-Mental Status Examination [23]; BI, the Barthel-100 index [24]; NOT-S, the Nordic Orofacial Test-Screening [25]; WT, the Water Swallow Test [26]; WHO-5, the WHO-five Well-Being index [34]; TUG, the Timed Up & Go test [36]; SD, standard deviation.

significant (Table II, right columns). The MISA-DK total scale correlated strongly with the MMSE, the BI and the NOT-S; the positioning subscale correlated strongly with the BI; the self-feeding skills subscale correlated strongly with the MMSE and with the BI; the solid ingestion subscale correlated strongly with the MMSE, the BI, and the NOT-S; and the texture management-solids subscale correlated strongly with the MMSE. The liquid ingestion subscale and the texture management-liquids subscale correlated less strongly to the convergent variables.

The multivariate analysis revealed that for the MISA-DK total scale, 55% of the variance in ingestive skills was explained by three of the convergent variables (Table III). Cognition appeared to be the most important factor, which accounted for 40% of the variance. For the subscales, the total explained variance ranged from 21% to 48%. Cognition was the most important factor for texture management of solids and liquids, although the contribution only explained 32% and 18% of the variance respectively. Physical function was the most important factor on positioning, self-feeding skills, and liquid and solid ingestion.

Validation by known-groups showed statistical significant differences of all MISA-DK scales in terms of frailty and on four of the MISA-DK scales in terms of the presence of pneumonia (Table IV).

Internal construct validity

Step 1: The study sample was distributed into three class intervals. Initially, the MISA-DK deviated significantly from

Table II. Internal consistency reliability and convergent validity of the MISA-DK.

	Reliability	Correlation of MISA-DK scales and the convergent variables							
MISA-DK scales	Cronbach's α	MMSE	p	BI	p	WT	p	NOT-S	p
Positioning	0.79	0.43	<0.001	**0.70**	<0.001	0.10	0.319	–0.34	<0.001
Self-feeding skills	0.85	**0.54**	<0.001	**0.66**	<0.001	0.15	0.117	–0.45	<0.001
Liquid ingestion	0.83	0.42	<0.001	0.40	<0.001	0.39	<0.001	–0.39	<0.001
Solid ingestion	0.90	**0.51**	<0.001	**0.61**	<0.001	0.31	0.001	**–0.52**	<0.001
Texture management-solids	0.80	**0.50**	<0.001	0.41	<0.001	0.24	0.014	–0.39	<0.001
Texture management-liquids	0.77	0.41	<0.001	0.32	0.001	0.21	0.029	–0.25	0.010
MISA-DK total scale	0.95	**0.59**	<0.001	**0.66**	<0.001	0.31	0.001	**–0.53**	<0.001

MISA-DK, the Danish version of the McGill Ingestive Skills Assessment [11,16]; MMSE, the Mini-Mental Status Examination [23]; BI, the Barthel-100 index [24]; NOT-S, the Nordic Orofacial Test-Screening [25]; WT, the Water Swallow Test [26].
Hypothesised strong correlations are shaded, and bold highlights strong correlations, Spearman's rho (r_s) > 0.50.

Table III. Contribution of the convergent variables on ingestive skills measured with the MISA-DK.

Variables	F (df1,df2)	p	R^2 change	Standardised β	p
Positioning (total explained variance 45%)	83.3 (1,100)	<0.001			
BI			0.45	0.674	<0.001
Self-feeding skills (total explained variance 48%)	44.9 (2,99)	<0.001			
MMSE			0.06	0.293	0.001
BI			0.42	0.476	<0.001
Liquid ingestion (total explained variance 37%)	18.9 (3,98)	<0.001			
MMSE			0.04	0.235	0.021
BI			0.22	0.307	0.003
WT			0.11	0.312	<0.001
Solid ingestion (total explained variance 48%)	30.3 (3,98)	<0.001			
MMSE			0.08	0.304	0.001
BI			0.35	0.399	<0.001
WT			0.05	0.232	0.002
Texture management-solids (total explained variance 35%)	25.9 (2,99)	<0.001			
MMSE			0.32	0.472	<0.001
NOT-S			0.03	–0.196	0.038
Texture management-liquids (total explained variance 21%)	13.1 (2,99)	<0.001			
MMSE			0.18	0.394	<0.001
WT			0.03	0.181	0.048
MISA-DK total scale (total explained variance 55%)	40.7 (3,98)	<0.001			
MMSE			0.40	0.366	<0.001
BI			0.10	0.403	<0.001
WT			0.05	0.222	0.002

MISA-DK, the Danish version of the McGill Ingestive Skills Assessment [11,16]; MMSE, the Mini-Mental Status Examination [23]; BI, the Barthel-100 index [24]; NOT-S, the Nordic Orofacial Test-Screening [25]; WT, the Water Swallow Test [26].
Stepwise multiple regression analysis with MISA-DK scales as dependent variables and convergent variables as independent variables.

Table IV. Known-group validity of the MISA-DK.

		MISA-DK						
Characteristic	Statistic[a]	Positioning	Self-feeding	Liquid Ingestion	Solid ingestion	Texture management-solids	Texture management-liquids	MISA-DK total scale
Frailty								
Absent (n = 40)	Mean ± SD	10.8 ± 1.6	19.2 ± 2.0	18.8 ± 1.9	31.8 ± 4.3	19.9 ± 3.3	13.6 ± 2.4	114.1 ± 11.0
Present (n = 64)	Mean ± SD	8.6 ± 1.9	16.0 ± 3. 5	16.5 ± 3.3	26.2 ± 5.2	16.6 ± 4.6	12.1 ± 2.8	96.0 ± 17.2
	Z	−5.912	−5.266	−3.595	−5.414	−3.647	−2.874	−5.749
	p value	<0.001	<0.001	<0.001	<0.001	<0.001	0.004	<0.001
Pneumonia								
Absent (n = 62)	Mean ± SD	9.4 ± 2.1	17.6 ± 3.2	18.1 ± 2.8	29.3 ± 5.1	18.8 ± 3.8	13.1 ± 2.5	106.2 ± 15.6
Present (n = 48)	Mean ± SD	9.3 ± 2.0	16.8 ± 3.4	16.7 ± 3.2	27.1 ± 5.8	16.6 ± 4.7	12.2 ± 2.8	98.7 ± 18.2
	Z	−0.317	−1.258	−2.178	−1.984	−2.700	−1.697	−2.116
	p value	0.753	0.209	0.029	0.047	0.007	0.090	0.034

MISA-DK, the Danish version of the McGill Ingestive Skills Assessment [11,16].
[a]Mann Whitney U-test and a two-sided significance level of 0.05.

Table V. Rasch analysis of MISA-DK: summary of fit statistics.

	Item-person interaction		Item-trait interaction		Reliability	Unidimensionality t-test (%) (95% CI when % >5.0)
Analysis[a]	Item residual Mean (SD)	Person residual Mean (SD)	Chi-square χ^2 (df)	p value	PSI	
1)	−0.02 (1.81)	0.10 (1.18)	313 (86)	<0.001	0.93	21.8 (17.7; 25.9)
2)	−0.13 (1.51)	−0.08 (1.10)	182 (86)	<0.001	0.94	16.4 (12.3; 20.4)
3)	−0.09 (1.13)	−0.11 (1.06)	131 (74)	<0.001	0.93	18.2 (14.1; 22.3)
4)	−0.29 (1.35)	−0.35 (0.94)	12 (12)	0.424	0.85	4.6 (0.5; 8.6)
Satisfactory fit	0 (< 1.4)	0 (< 1.4)		>0.050	>0.7	<5% or lower CI <5%

MISA-DK, the Danish version of the McGill Ingestive Skills Assessment [11,16]; SD, standard deviation; df, degrees of freedom; PSI, Person Separation Index; CI, confidence interval.
[a]1) Step 1, initial analysis; 2) Step 2, resolving disordered thresholds; 3) Step 3, deletion of misfitting items; 4) Step 4, creation of six testlets.

the Rasch model with an item fit residual SD of 1.81 and a significant item-trait interaction (χ^2 (df) = 313 (86), p < 0.001) (Table V, analysis 1).

The person fit residual was satisfactory indicating no serious misfit among patients in the sample. The person parameters indicated no extreme scores and all but two patients fell within a fit residual range of ±2.5. The PSI of 0.93 demonstrated excellent reliability. A lack of unidimensionality was evident with 21.8% statistical significant different person estimates based on two item subsets.

At individual item level, misfit was found for six items (Table VI, step 1). Disordered thresholds were found for 11 items (items 7,12,13,32,34,35,37,38,39,42,43). No DIF by gender or age was displayed. Local dependency was identified by residual correlations >0.2 for several item pairs within each of the six subscales.

Step 2: The threshold disordering for the 11 items involved the score 2, which did not have a range along the ability scale where it was the most likely category. This was resolved by combining score 1 and score 2 for these items. Model fit improved slightly, but not satisfactorily and a lack of unidimensionality was evident (Table V, analysis 2). At individual item level, misfit was resolved for three items, but three additional items displayed misfit (Table VI, step 2). Indication of local dependency was persistent.

Step 3: The misfitting items were removed. This provided satisfactory item-person interaction fit statistics. The item-trait interaction (χ^2 (df) = 131 (74), p < 0.001) still indicated model misfit (Table V, analysis 3). Multidimensionality and local dependency were persistent.

Step 4: As step 3 did not provide model fit, further analysis was undertaken of the MISA-DK scale from step 2. The residual correlations within each of the six subscales displayed a pattern, and the items were grouped together into six testlets. Model fit and the test for unidimensionality became satisfactory (Table V, analysis 4). The PSI decreased to 0.85. Additional reliability testing using Cronbach's α resulted in 0.88. Testlet 4, which corresponds to the solid ingestion subscale, displayed nonsignificant misfit (Table VI, step 4), but the item characteristic curve displayed model fit (Figure 1).

No DIF by gender or age was displayed and no further local dependency was observed. The MISA-DK appeared reasonable targeted to the sample (Figure 2). The item location mean (SD) was 0.0 (0.221), whereas the person location mean (SD) was 0.537 (0.612), which may indicate that this sample on average was of a higher ability level than the average of the scale. Seven percent of the sample was not covered by the scale.

Discussion

We examined the validity of the MISA-DK using classical test theory and the Rasch model. From a classical test theory perspective, we found support for the internal consistency

Table VI. Items of MISA-DK demonstrating misfit to the Rasch model in analysis step 1 and 2 and item (testlet) fit in step 4.

Item number and item descriptor	Location[a]	SE[b]	Residual[c]	χ^{2d}	(df)[e]	p	F-Statistic[f]	(df1,df2)[e]	p
Step 1									
9. Able to focus on meal	−0.34	0.18	**3.26**	7.5	(2)	0.024	3.0	(2,107)	0.055
31. Capable of eating heterogeneous textures	0.27	0.17	**−3.34**	10.3	(2)	0.006	**10.0**	(2,107)	**<0.001**
35. Capable of eating sticky solids	0.47	0.15	**2.68**	6.5	(2)	0.039	2.1	(2,107)	0.129
37. Capable of eating puree	0.39	0.15	**4.07**	**21.1**	(2)	**<0.001**	4.7	(2,107)	0.011
38. Capable of eating pudding	1.10	0.13	**5.09**	**134.0**	(2)	**<0.001**	**31.2**	(2,107)	**<0.001**
42. Capable of drinking honey consistency	0.08	0.16	**2.78**	**19.1**	(2)	**<0.001**	5.6	(2,107)	0.005
Step 2									
9. Able to focus on meal	−0.45	0.19	**3.82**	**16.7**	(2)	**<0.001**	5.6	(2,107)	0.005
16. Able to take a sequence of sips	1.26	0.16	**2.85**	5.0	(2)	0.081	2.7	(2,107)	0.071
31. Capable of eating heterogeneous textures	0.14	0.18	**−3.20**	5.9	(2)	0.053	5.9	(2,107)	0.004
36. Capable of eating soft solids	−0.30	0.18	**3.19**	1.3	(2)	0.524	0.6	(2,107)	0.545
38. Capable of eating pudding	1.78	0.23	2.37	**33.9**	(2)	**<0.001**	**12.5**	(2,107)	**<0.001**
39. Capable of drinking water	−0.01	0.25	−1.78	11.4	(2)	0.003	11.0	(2,107)	**<0.001**
Step 4									
1. Positioning scale	−0.09	0.06	0.91	5.82	(2)	0.055	3.60	(2,107)	0.031
2. Self-feeding skills scale	0.05	0.05	−0.19	1.20	(2)	0.549	0.27	(2,107)	0.761
3. Liquid ingestion scale	−0.13	0.05	−0.99	1.48	(2)	0.478	1.15	(2,107)	0.320
4. Solid ingestion scale	0.08	0.03	**−2.60**	1.00	(2)	0.607	0.86	(2,107)	0.426
5. Texture management scale-solids	0.37	0.05	1.04	1.35	(2)	0.508	0.82	(2,107)	0.445
6. Texture management scale-liquids	−0.27	0.06	0.07	1.40	(2)	0.489	0.82	(2,107)	0.445

MISA-DK, the Danish version of the McGill Ingestive Skills Assessment [11,16].

[a]Expressed in logits. Positive values reflect difficult items and negative values reflect easy items; [b]SE, standard error; [c]Residuals summarise the deviation of observed from expected responses. Values outside the range of ±2.5 indicates misfit and are bold; [d]χ^2 (chi-square values) summarise the deviation of observed from expected responses across three class intervals of patients. Higher values represent larger deviations. Bonferroni adjusted statistically significant deviations (p value of 0.001 in step 1 & 2 and p value of 0.008 in step 4) indicate misfit and are bold; [e]df, degrees of freedom; [f]F-statistics from one-way ANOVA of deviations from model expectations across the three class intervals of patients. Bonferroni adjusted statistically significant deviations (p value of 0.001 in step 1 & 2 and p value of 0.008 in step 4) indicate misfit and are bold.

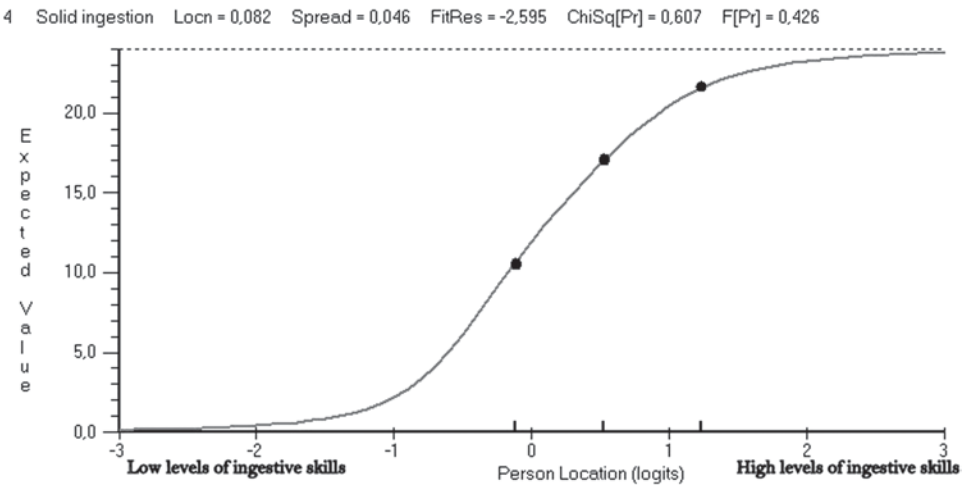


Figure 1. The item characteristic curve (ICC) for testlet 4 (solid ingestion subscale of the Danish version of the McGill Ingestive Skills Assessment). The three dots shown along the curve are the observed means of patients distributed into three class intervals [38]. The observed means for testlet 4 follow the ICC, which implies that the testlet is functioning consistently across the three class intervals.

reliability of the MISA-DK scales, which is consistent with the Canadian validation study of the MISA [15]. Though, we found a higher Cronbach's α of 0.95 for the MISA-DK total scale. A Cronbach's α > 0.90 may indicate item redundancy [17] or multidimensionality [20]. Our findings may be due to the translation and possible semantic nonequivalence of the MISA-DK. However, classical psychometric methods are dependent on the sample [20]. As the Canadian sample was primarily recruited from long term-care facilities and ours was recruited from acute care, there is likely a difference in the samples.

The results of the convergent validation supported partly our expectations. For the MISA-DK total scale, we found strong correlations with the MMSE, BI and NOT-S, but only fair correlation with the WT. The results of the correlations to physical function and cognition are consistent with Lambert et al. [15]. However, we found a higher magnitude of the correlation coefficients, which may be a reflection of the sample dependency using classical psychometric methods [20]. Lambert et al. [15] used the Modified MMSE [46], which includes the ability to give personal information. In our study, patients who were not able to give such information were

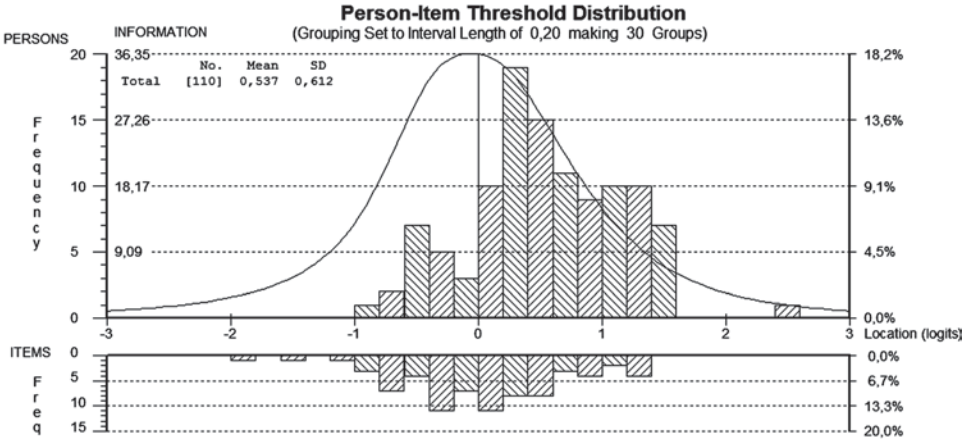


Figure 2. Targeting map of the Danish version of the McGill Ingestive Skills Assessment after rescoring and creating six testlets. The testlet-threshold location range from −2 to about 1.5 logits, and person locations range from −1 to about 2.5 logits. The full range of person locations in the study sample is not covered by the scale.

excluded. Therefore, our sample might have been at a higher cognitive ability level than the Canadian sample. Additionally, Lambert et al. [15] included age and not swallowing and orofacial function as convergent variables. In line with our findings in the preparation for the multivariate analyses, Lambert et al. [15] did not find strong correlation between MISA and age. Overall, the convergent variables explained 55% of the variance in ingestive skills measured by the MISA-DK total scale, and cognitive and physical function accounted for 50%. This may reflect the importance of acknowledging a comprehensive approach when measuring swallow efficacy and safety in dysphagic patients [5,6,12].

For the subscales of the MISA-DK, we found the expected correlations for positioning and self-feeding skills. The NOT-S correlated strongly with solid ingestion, but only fairly with liquid ingesting and texture management, which was not expected. A possible explanation may be that the physiology of eating and drinking are not identical [6]. It cannot be excluded that the domains in the NOT-S reflect the skills needed for eating more than the skills needed for drinking. Surprisingly, the WT did not demonstrate strong correlations to any of the subscales covering oral motor and pharyngeal skills for liquids and solids and texture management. When we applied multivariate analysis, swallowing function turned out to have a small impact on ingestion of liquids and solids and management of liquids, whereas orofacial functions did not. Although isolated to the texture of water, the WT might be based on activity performance compared to the NOT-S, which assesses oral motor function separated from an activity. It is recognised that the relationships between motor impairments and activity limitations are not straightforward [47]. It is worth noticing that the WT is recognised to display high sensitivity and low specificity [48]. As 71% failed the WT in our sample, it cannot be excluded that some of the patients could be false positive. Yet, it may remain a hypothesis as the sensitivity and specificity of the MISA-DK is not established.

Using bivariate and multivariate analysis, the two texture management scales were least explained by the convergent variables, and we found an unexplained variance of 68% and 82% respectively, which is in line with Lambert et al. [15]. This raises questions regarding what these scales are measuring.

In support of the validity of the MISA-DK, all of the scales were able to discriminate effectively between known groups. Patients categorised as frail were rated to have lesser ingestive skills versus patients who were not frail, and the presence of pneumonia was related to a decrease in liquid ingestion, solid ingestion and texture management of solids. This is in concordance with current research [1,2,5].

From the Rasch model perspective, the internal construct validity of the MISA-DK was not initially supported. We found some problems with disordered thresholds, misfitting items, local dependency and signs of multidimensionality.

Threshold disordering was present for 11 items, of which seven items belonged to the texture management scales, and rescoring did not provide overall model fit. The observed disordering among the score categories may indicate that the score scale does not work as intended. Disordered thresholds may occur if the labeling of options is similar to one another, potentially confusing or open to misinterpretation [21]. This might refer back to the translation and possible semantic non-equivalence of the MISA-DK, or may reflect the uncertainty of what the texture management scales are measuring. For the misfitting items, a majority also belonged to the texture management scales. Indication of multidimensionality and redundancy was displayed for these misfitting items. This finding may also clarify the results of our analysis of the convergent validity for these two scales. Therefore, the same problem needs to be identified in a larger sample with Danish and Canadian groups to indicate whether the existing response category structure and the misfitting items should be reconsidered or the Danish translation should be revised.

From our results, it seems that the main problem was a local dependency problem. Creating six testlets of the subscales and treating them as six separate items provided evidence of model fit and unidimensionality of the MISA-DK. The decrease in the reliability indices may reflect an influence of local dependency of the items within each subscale [29], and may confirm the former discussion of the high internal

consistency reliability of 0.95. Whether the identified local dependency is response or trait dependence may be difficult to distinguish in polytomous analysis [29]. Therefore, further analysis of the MISA-DK total scale and each subscale is needed. Additionally, if a shorter version of the MISA-DK is sought, then our results might be taken into account in the item selection. In its present form, it seems that the local dependency can be accounted for *post hoc* with acceptable levels of reliability above the minimum requirement of 0.70. Besides, our results confirm that no DIF by age or gender was present. This indicates that the probability of being rated on a particular score is not dependent of such external factors. However, it could be beneficial to analyse for other factors such as diagnoses.

Some methodological issues related to this study have to be considered. Although the research assistant (RA) was trained in the measurement of the convergent variables, it cannot be excluded that there have been differences in the severity of the judgements between RA and TH. Additionally, only one rater performed the MISA-DK. Thus, the inter- and intra-rater reliability of the MISA-DK still remains to be established. For the convergent validity, it could be argued that the criteria for strong correlation were too high when dealing with a complex construct as ingestion. Because there are no widely accepted criteria for defining a strong correlation [27], we applied multivariate analysis to form a better understanding of the construct of the MISA-DK. When using the BI, MMSE, NOT-S and the WT as convergent variables, an unexplained variance of 45% was evident. Some studies have raised concerns about the dimensionality of BI [49] and MMSE [50] as well as about the diagnostic precision of the WT [48]. This may suggest inclusion of other well validated convergent variables for further validation of the MISA-DK. The operational definitions of the frailty criteria differed from Fried et al. [32] in terms of exhaustion, which we measured using the WHO-5 and the reduced physical activity which we measured by a BI score <50. However, comparable modifications have been implemented in other studies [51].

This is the first study addressing the dimensionality of the MISA-DK. In the Rasch analysis, we applied Bonferroni adjusted χ^2 and F-statistics. Analysis without adjustments could have been applied and might have resulted in a larger percentage of significant χ^2 and F-statistics, indicating item misfit [41]. However, as it seemed that local dependency was the main problem, we decided not to do so. Our study sample was on average of a higher ability level than the average of the scale, which is displayed in Figure 1. The sample tends to have relative high values on the trait relative to the origin of 0.0 of the items, which do map a continuum from less to more. This means that overall our sample tend not to show the low levels of ingestive skills covered by the items in the MISA-DK. The exclusion criteria used in this study may have caused this. A sample at lower levels of ingestive skills may reduce the slight tendency of a ceiling effect. Additionally, the sample size was relative low and distributed into three class intervals. *Post hoc* analysis with two class intervals did not deviate from the present results. Finally, as the sample size is relative low, we have not provided an exchange rate between the raw score and the Rasch transformed scores.

Conclusion

The results of this study indicate that the internal consistency and external construct validity of the MISA-DK equal the original Canadian version of the MISA. Thus, measurement equivalency is established. However, we found some indication of multidimensionality in the MISA-DK scale, which could be explained by local dependency. Although achieving good fit to the Rasch model after adjustments, additional studies are needed to establish cross-cultural validity. Like this, it is possible to verify whether the existing response category structure should be reconsidered, and whether reduction of the items in the MISA and MISA-DK is necessary. Finally, establishment of inter- and intra-rater reliability of the MISA-DK is needed.

Acknowledgements

We are grateful to all the participating patients and the facilities' staff, and we thanks occupational therapist Charlotte Ehlers Hansen for throughout assistance in the data collection.

***Declaration of interest*:** The study was financial supported by the Research council at Herlev University Hospital, the Research Foundation of the Danish Occupational Therapy Association (FF2/09-1) and the Lundbeck Foundation (FP03/2011). The authors alone are responsible for the content and writing of the paper. No party supporting this article has or will confer a benefit on us or on any organization with which we are associated.

References

1. van der Maarel-Wierink CD, Vanobbergen JN, Bronkhorst EM, Schols JM, de Baat C. Risk factors for aspiration pneumonia in frail older people: a systematic literature review. J Am Med Dir Assoc 2011;12:344–354.
2. Cabre M, Serra-Prat M, Palomera E, Almirall J, Pallares R, Clavé P. Prevalence and prognostic implications of dysphagia in elderly patients with pneumonia. Age Ageing 2010;39:39–45.
3. Miller N, Carding P. Dysphagia: implications for older people. Rev Clin Gerontol 2007;17:177–190.
4. Bergman H, Ferrucci L, Guralnik J, Hogan DB, Hummel S, Karunananthan S, Wolfson C. Frailty: an emerging research and clinical paradigm–issues and controversies. J Gerontol A Biol Sci Med Sci 2007;62:731–737.
5. Rofes L, Arreola V, Romea M, Palomera E, Almirall J, Cabré M, Serra-Prat M, Clavé P. Pathophysiology of oropharyngeal dysphagia in the frail elderly. Neurogastroenterol Motil 2010;22:851–8, e230.
6. Cichero JYA, Murdoch BE. Dysphagia: foundation, theory and practice. Chichester: Wiley; 2006.
7. Clark GF, Avery-Smith W, Wold LS, Anthony P, Holm SE. Eating and Feeding Task Force; Commission on Practice. Specialized knowledge and skills in feeding, eating, and swallowing for occupational therapy practise. Am J Occup Ther 2007;61:686–699.
8. Fisher AG. Occupational therapy intervention process model: a model for planning and implementing top-down, client-centred, and occupation-based interventions. Fort Collins, CO: Three Star Press; 2006.
9. Laver Fawcett A. Principles of assessment and outcome measurement for occupational therapists and physiotherapists: theory, skills and application. Chichester: John Wiley and Sons Ltd; 2007.

10. Hansen T, Kjaersgaard A, Faber J. Measuring elderly dysphagic patients' performance in eating - a review. Disabil Rehabil; 1–10. Early online, 2011. Available at: http://informahealthcare.com/doi:10.3109/09638288.2011.553706. Accessed February 2011.
11. Lambert HC, Gisel EG, Wood-Dauphine S, Groher ME, Abrahamowicz M. McGill Ingestive Skills Assessment: user's manual. Ottawa: Canadian Association of Occupational Therapists; 2006.
12. Leopold NA, Kagel MC. Dysphagia–ingestion or deglutition?: a proposed paradigm. Dysphagia 1997;12:202–206.
13. Lambert HC, Gisel EG, Groher ME, Wood-Dauphinee S. McGill Ingestive Skills Assessment (MISA): development and first field test of an evaluation of functional ingestive skills of elderly persons. Dysphagia 2003;18:101–113.
14. Lambert HC, Abrahamowicz M, Groher M, Wood-Dauphinee S, Gisel EG. The McGill ingestive skills assessment predicts time to death in an elderly population with neurogenic dysphagia: preliminary evidence. Dysphagia 2005;20:123–132.
15. Lambert HC, Gisel EG, Groher ME, Abrahamowicz M, Wood-Dauphinee S. Psychometric testing of the McGill Ingestive Skills Assessment. Am J Occup Ther 2006;60:409–419.
16. Hansen T, Lamberts HC, Faber J. Content validation of a Danish version of "The McGill Ingestive Skills Assessment" for dysphagia management. Scand J Occup Ther; 1–12. Early online 2010. Available at: http://informahealthcare.com/doi:10.3109/11038128.2010.521949. Accessed October 2010.
17. Streiner D, Norman G. Health measurement scales. 3rd ed. Oxford: Oxford University Press; 2003.
18. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum 2007;57:1358–1362.
19. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989;70:857–860.
20. Crocker LM, Algina J. Introduction to classical and modern test theory. Mason, OH: Wadsworth Group/Thomas Learning; 2006.
21. Bond TG, Fox CM. Applying the Rasch model. Fundamental measurement in the human sciences. London: Lawrence Erlbaum Associates; 2001.
22. Wilson M. Constructing measures: an item response modeling approach. Mahwah, NJ: Erlbaum; 2005.
23. 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.
24. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989;42:703–709.
25. Bakke M, Bergendal B, McAllister A, Sjögreen L, Asten P. Development and evaluation of a comprehensive screening for orofacial dysfunction. Swed Dent J 2007;31:75–84.
26. Smithard DG, O'Neill PA, Park C, England R, Renwick DS, Wyatt R, Morris J, Martin DF; North West Dysphagia Group. Can bedside assessment reliably exclude aspiration following acute stroke? Age Ageing 1998;27:99–106.
27. Portney IG, Watkins MP. Foundations of clinical research. Application to practice. 3rd ed. Upper Saddle River, NJ: Prentice Hall Health; 2009.
28. Masters G. A Rasch model for partial credit scoring. Psychometrika 1982;47:149–174.
29. Marais I, Andrich D. Formalizing dimension and response violations of local independence in the unidimensional Rasch model. J Appl Meas 2008;9:200–215.
30. Kjaersgaard A. Ansigt, mund og svælg: Undersoegelse og behandling efter Coombes konceptet. (Face, mouth and throat: assessment and treatment a.m. the Coombes concept). Copenhagen: FADL; 2005.
31. Kørner EA, Lauritzen L, Nilsson FM, Wang A, Christensen P, Lolk A. Mini mental state examination. Validation of new Danish version. Ugeskr Laeg 2008;170:745–749.
32. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, et al.; Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56:M146–M156.
33. Kondrup J, Rasmussen HH, Hamberg O, Stanga Z; Ad Hoc ESPEN Working Group. Nutritional risk screening (NRS 2002): a new method based on an analysis of controlled clinical trials. Clin Nutr 2003;22:321–336.
34. Heun R, Burkart M, Maier W, Bech P. Internal and external validity of the WHO Well-Being Scale in the elderly general population. Acta Psychiatr Scand 1999;99:171–178.
35. Bohannon RW, Peolsson A, Massey-Westropp N, Desrosiers J, Bear-Lehman J. Reference values for adult grip strength measured with a Jamar dynamometer: a descriptive meta-analysis. Physiotherapy 2006;92:11–15.
36. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39:142–148.
37. Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). Br J Clin Psychol 2007;46:1–18.
38. Andrich D, Sheridan B, Luo G. Interpreting RUMM2030. Perth: RUMM Laboratory Pty; 2009.
39. Smith EV Jr. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. J Appl Meas 2002;3:205–231.
40. Fisher WP. Reliability Statistics. Rasch Measure Trans 1992;6:238.
41. Tennant A, Pallant JF. Unidimensionality matters! (a tale of two Smiths?). Rasch Measure Trans 2006;20:1048–1051.
42. Guemin L, Robert LB, David AF. Incorporating the testlet concept in test score analyses. Educ Meas I P 2000;19:9–15.
43. Linacre JM. Optimizing rating scale category effectiveness. J Appl Meas 2002;3:85–106.
44. Altman DG. Practical statistics for medical research. CRC Press; 1990.
45. Linacre JM. Sample size and item calibration stability. Rasch Measur Trans 1994;7:328.
46. Tombaugh T, McDowell I, Kristjansson B. Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): a psychometric comparison and normative data. Psychol Assess 1996;8:48–59.
47. Vandervelde L, Van den Bergh PY, Renders A, Goemans N, Thonnard JL. Relationships between motor impairments and activity limitations in patients with neuromuscular disorders. J Neurol Neurosurg Psychiatr 2009;80:326–332.
48. Bours GJ, Speyer R, Lemmens J, Limburg M, de Wit R. Bedside screening tests vs. videofluoroscopy or fibreoptic endoscopic evaluation of swallowing to detect dysphagia in patients with neurological disorders: systematic review. J Adv Nurs 2009;65:477–493.
49. de Morton NA, Keating JL, Davidson M. Rasch analysis of the barthel index in the assessment of hospitalized older patients after admission for an acute medical condition. Arch Phys Med Rehabil 2008;89:641–647.
50. Schultz-Larsen K, Kreiner S, Lomholt RK. Mini-Mental Status Examination: mixed Rasch model item analysis derived two different cognitive dimensions of the MMSE. J Clin Epidemiol 2007;60:268–279.
51. Romero-Ortuno R, Walsh CD, Lawlor BA, Kenny RA. A frailty instrument for primary care: findings from the Survey of Health, Ageing and Retirement in Europe (SHARE). BMC Geriatr 2010;10:57.

Paper III

Scandinavian Journal of Occupational Therapy. 2012; 19: 488–496

informa
healthcare

ORIGINAL ARTICLE

Reliability of the Danish version of the McGill Ingestive Skills Assessment for observation-based measures during meals

TINA HANSEN[1], HEATHER C. LAMBERT[2] & JENS FABER[3,4]

[1]*Department of Occupational Therapy, Herlev University Hospital, Herlev, Denmark,* [2]*School of Physical and Occupational Therapy, McGill University, Montreal, Canada,* [3]*Department of Medicine/Endocrinology, Herlev University Hospital, Herlev, Denmark, and* [4]*University of Copenhagen, Denmark*

Abstract

Aim: To establish measurement equivalence in terms of reliability of the Danish version of the Canadian McGill Ingestive Skills Assessment (MISA) for use by occupational therapists. *Methods:* A cross-sectional two-rater and test–retest design was applied. A total of 102 elderly medical patients were included consecutively, and were video-recorded during a meal. Raters were paired randomly for each video-case, which was re-scored within three to eight weeks. Reliability was evaluated with the intra-class correlation coefficients (ICC), the standard error of measurement (SEM), the smallest detectable change (SDC), and limits of agreement (LOA). *Results:* Inter-rater reliability was good to excellent (ICC $_{1.1}$ 0.61–0.84) and intra-rater reliability was excellent (ICC $_{3.1}$ 0.84–0.93). For the total scale, SEM was 7% between raters and 4% in repeated measurement by the same rater. For the absolute total scale range on 86 points, the SDC was 15.8 between raters and 10.3 in repeated measurement by the same rater. *Conclusions:* The reliability of the Danish MISA equals the original version and is suitable for clinical practice. When extending the evaluation of the reproducibility, weaker precision was evident when measurements are repeated by different raters than by the same rater. Therefore further investigation of rater effects is recommended.

Key words: *reproducibility of results, outcome assessment (health care), occupational therapy, geriatric, eating and drinking*

Introduction

Dysphagia is prevalent among frail elders (1,2), and may impair the ability to maintain quality in task performance while eating and drinking and/or to maintain normative expectations for appropriate mealtime behavior (3,4).This may lead to social isolation and reduced quality of life (5).

In order to provide adequate management, dysphagia requires careful and comprehensive examination (6). Dysphagia is currently assessed using clinical bedside assessments which includes anamnesis, evaluation of oral, pharyngeal, and laryngeal sensory and motor function, and water swallow tests (7). Also instrumental techniques such as videofluoroscopic or fiberoptic endoscopic examination of swallowing (7) are used. However, these assessments are often performed in an artificial environment (6,7) and may not fully reflect the complexity of eating and drinking in a natural context (6). Therefore, it would be beneficial to complement these methods with information on the dysphagic patient's task performance during a natural meal (6,7).

In multidisciplinary dysphagia management, occupational therapists consider the interplay of physical, cognitive, environmental, and sociocultural factors in order to assist the dysphagic patient to return to efficient and safe performance in eating and drinking activities (8). A recent review of the international literature on evidence-based assessment tools measuring dysphagic elders' performance during a natural meal revealed one occupational therapy assessment tool with satisfactory psychometric properties (9): the McGill Ingestive Skills Assessment

Correspondence: Tina Hansen, Occupational Therapist, MSc.OT, Department of Occupational Therapy 53P1, Herlev University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark. Tel: +45 24 85 35 86/ +45 44 88 38 18. E-mail: ergotp.tina@gmail.com

(Received 30 August 2011; revised 25 February 2012; accepted 6 March 2012)

ISSN 1103-8128 print/ISSN 1651-2014 online © 2012 Informa Healthcare
DOI: 10.3109/11038128.2012.674552

(MISA) (10). The MISA, which was developed in Canada, is a relatively new method for measuring the ability of frail elders to eat and drink safely and independently during a natural meal (11). The MISA is intended for use in diagnosis, treatment planning, and evaluation (10).

The conceptualization of eating and drinking in the MISA is based on a construct termed "Ingestion" (10,12). Ingestion includes cognition, physiological factors such as hunger, exteroceptive sensation of the meal, neck and truncal position, the manual activities and oromandibular aspects of eating and drinking as well as the voluntary, automatic, and reflex components of bolus preparation and the swallow (12). Thus, ingestion includes the actions of self-feeding (i.e. the process of setting up, arranging, and bringing food/liquid from the plate/cup to the mouth), eating (i.e. the ability to keep and manipulate food/liquid in the mouth and swallow it) and swallowing (i.e. the complicated act where food, liquid, medication, or saliva is moved from the mouth through the pharynx and esophagus into the stomach (8,13).

The items in the MISA have been generated and psychometrically tested using classical test theory. The total scale has high inter-and intra-rater reliability (intra-class correlation coefficient (ICC) of 0.85 and 0.96), test–retest reliability (ICC of 0.88) (14) and high internal consistency (Cronbach's α above 0.70) (10). In a geriatric population, it correlates with constructs related to ingestion in terms of cognitive function and functional ability, it discriminates among groups with different levels of disabilities in terms of health status and denture wear (14), and it predicts time to death and to pulmonary infection (15).

As the MISA provides evidence-based measurements not just for diagnosis but also for treatment planning and evaluation, it has been considered to be of value for Danish occupational therapists and has been translated into Danish (MISA-DK) (16). Translation and adaptation of assessment instruments involves equivalence (17). Conceptual and semantic equivalence of the MISA-DK has been addressed through a comprehensive translation procedure, expert panel judgments, and pilot testing (16). Measurement equivalence in terms of construct validity and internal consistency has also been established (18). The MISA-DK has high internal consistency (Cronbach's α above 0.70), correlates strongly with measures of cognitive function, functional ability, and orofacial function, discriminates among groups in terms of frailty status and pneumonia, and demonstrates one single construct in Rasch analysis (18). However, measurement equivalence in terms of reliability of the MISA-DK remains to be established. As the MISA is an observation-based assessment measuring dysphagic elders' performance during a mealtime, both inter- and intra-rater reliabilities need to be addressed (19,20).

Reliability refers to the reproducibility of measurements, which concerns the degree to which repeated measurements in stable study subjects provide similar results (19). The term reproducibility covers two concepts, relative and absolute reliability (21), which are often used interchangeably but are in fact conceptually distinct (20). Relative reliability is defined as the ratio of variability between subjects to the total variability of all measurements in the sample, and absolute reliability is the degree to which scores or ratings are identical (20–22). Relative reliability parameters are required for measurements that are used for discriminative purposes, and absolute reliability parameters are required for measurements that are used for evaluative purposes (22). Since the MISA is intended to be diagnostic and evaluative, both concepts are important. However, only relative reliability has been addressed for the original version of the MISA (14). Thus, the aims of this study were to establish the measurement equivalence of the MISA-DK in terms of relative reliability, and to extend the evaluation of the reproducibility of the MISA-DK in terms of absolute reliability.

Methods

Participants

Patients consecutively admitted to two departments of general medicine at an acute hospital in the Capital Region of Copenhagen between December 2009 and February 2011 were screened for inclusion within 48 h of admission. The patients were invited to participate in the study if they were over 65 years, were not terminally ill, would require more than two days of hospitalization and were able to give personal information and written informed consent. The patients were excluded if they did not fulfill five criteria for direct swallowing evaluation (23), namely the ability to: remain alert for at least 15 min, sit in a chair or bed in at least a 60° upright position, swallow saliva, cough voluntarily, and clear the throat twice. Of 439 eligible patients, 168 were unable to give personal information and written informed consent and 87 declined. Of the remaining 184 patients, 74 (40%) were unable to perform the five swallowing criteria. This resulted in the inclusion of 110 patients for the construct validation study of the MISA-DK (18), of which 102 patients agreed to be video-recorded during a meal for this reliability study. The study was approved by the local ethical committee in the Capital Region (Reg. No: H-C-2009-061) and the Danish Data Protection Authority (Reg. No: 2009-41-3719).

Instrumentation

MISA-DK. The Danish translation of the MISA was used (16). It is composed of 43 items distributed into six subscales: (i) positioning (four items) addressing the patient's ability to maintain a position that is safe for eating and drinking; (ii) self-feeding skills (seven items) addressing the patient's self-feeding skills, behavior, and judgment; (iii) liquid ingestion (seven items) addressing the patient's oral motor and pharyngeal skills for liquids; (iv) solid ingestion (12 items) addressing the patient's oral motor and pharyngeal skills for solids; (v) texture management-solids (eight items) addressing the patient's ability to manage eight solid food textures; and (vi) texture management–liquids (five items) addressing the patient's ability to manage five liquid textures. Each item is scored on a three-point ordinal scale, which is summarized to give subscale scores and a total score. Increasing scores indicate increasing ability levels in ingestive skills (10,16).

Demographics and functional performance. In order to specify the sample population (20), information on sex, age, main diagnostic categories, and functional performance is presented. Functional performance in activities of daily living (ADL) was measured using the Barthel-100 index (BI), which covers domains related to self-care (feeding, grooming, bathing, dressing, bowel and bladder care, and toilet use) and to mobility (ambulation, transfers, and stair climbing). The score ranges from 0 to100, and increasing scores indicate higher physical function (24). Cognitive function was measured using the Mini-Mental Status Examination (MMSE), which covers seven domains of cognition (temporal orientation, spatial orientation, immediate memory, attention and calculation, recall, language, and visual construction). The score ranges from 0 to 30, and increasing scores indicate higher cognitive ability (25). Orofacial function was measured using the Nordic Orofacial Test-Screening (NOT-S) which contains a clinical examination with six domains (the face at rest, nose breathing, facial expression, masticatory muscle and jaw function, oral motor function, and speech). The score ranges from 0 to 6, and higher score indicates orofacial dysfunction (26). Swallowing function was measured with the water swallow test (WST) which included two stages. In stage 1, a teaspoon (5 ml) of water was given three times and those patients safe on at least two of three attempts were given a larger volume (60 ml) of water to drink continuously from a cup. The criteria for safe completion of stages 1 and 2 were: no delay or absence of up and forward laryngeal movement on attempted swallow, no cough or choking during or after the swallow, no change in voice quality, and no signs of respiratory distress. Failure at either stage was recorded as a failed WST (27).

Raters

Thirty-eight occupational therapists from 12 different acute and rehabilitation sites were recruited as raters. The raters' average length of time since graduation in occupational therapy was 7.1 ± 6.6 years (range 0.5–29 years) and the average of clinical experience in dysphagia management was 4.0 ± 4.1 years (range 0.5–17 years). Seventeen raters (44.7%) had specialized postgraduate education in dysphagia. The raters underwent an eight-hour training course given by the first author (TH), who is a senior occupational therapist with specialized knowledge and skills in dysphagia management. The training course involved four to 10 raters at a time. The course program included review of basic anatomy and physiology of eating, drinking, and swallowing, introduction to the MISA-DK, and practice of mealtime observation and scoring using digitized real-life examples of five patients. After each video case, the raters discussed their scorings. Conflicting viewpoints were resolved using the instruction manual and via feedback by TH. After the training course, the raters administered the MISA-DK to at least five patients in their own clinical settings. During this period, the raters had the opportunity to discuss their ratings with each other and with TH. To ensure that the training of the raters reached a pre-defined criterion regarding accuracy of rating, the raters rated a video case that was also rated by TH and the second author (HCL). In the case of substantial deviation from the criterion ratings, the rater received extra supervision before participation in the study.

Procedure

In order to ensure independence between raters and stability in the participants' performance, each of the participants were video-recorded during a mealtime, either at the bedside or in a special eating area on the ward. The meal contained all 13 food textures assessed in the MISA, and the participant received the same assistance from TH as would normally be given. The video-recordings were made from the time the meal was served until the participant had completed the meal, or until TH terminated the evaluation because the mealtime was judged to be dangerous to the participant. The video camera was placed so that the participant's head and trunk were kept in the frame at all times, and videos were taken at an angle of 45 degree so that postural and orofacial characteristics could be assessed. The video-recordings were saved onto a CD in mpeg format and lasted on average 24 min (range 8–43 min).

Table I. Sample demographic and functional performance ($n = 102$).

Sex, *n* (%)	
Men	52 (51%)
Women	50 (49%)
Mean age in years ± SD	81.9 ± 7.6
Main diagnostic category, *n* (%)[a]	
Circulatory	67 (66%)
Sequelae after stroke	25 (25%)
Respiratory	59 (58%)
Musculoskeletal system	35 (34%)
Diabetes mellitus	23 (23%)
Nervous system	11 (11%)
Functional performance	
BI, Mean ± SD	50.0 ± 31.9
MMSE, Mean ± SD	22.0 ± 5.4
NOT-S, Mean ± SD	2.81 ± 1.5
WST failed, *n* (%)	68 (66.7%)

Note: [a]An individual patient may have more than one diagnosis.

The video cases, MISA-DK score-sheets with basic demographic and diagnostic information about the participant, and information on the mealtime menu were personally handed over to the rater by TH. For the inter-rater reliability, the raters were paired randomly across the clinical settings in a two-rater design for each video case. Each rater scored on average five video cases (range 2–11). For the intra-rater reliability, the rater re-scored the same video-cases in a test–retest design within a time frame of three to eight weeks.

The measures on the functional performance of the participants were performed by a research assistant, who is an experienced occupational therapist. The BI was routinely completed by the facility nursing staff or by interview with the patient.

Data analysis

We performed all analysis using SPSS 17.0. Descriptive statistics were used to describe the demographic and functional performance profile of the sample.

We assessed the relative reliability of the MISA-DK by calculating the ICC for subscale sum-scores and the total sum-score. The ICC is based on analysis of variance (ANOVA) (19). For the inter-rater reliability, we applied model 1 ($ICC_{1.1}$), which is a one-way random effect model with raters as random effects (19). For the intra-rater reliability, we applied model 3 ($ICC_{3.1}$), which is a two-way mixed model with rater as fixed effect and subjects as random effects (19). For the purpose of analysis, ICCs > 0 0.75 indicated excellent reliability, ICCs between 0.60 and 0.74 indicated good reliability, ICCs between 0.40 and 0.59 indicated fair reliability, and ICCs < 0.40 indicated poor reliability (28).

We assessed the absolute reliability of subscale sum-scores and total sum-score of the MISA-DK by calculating the standard error of measurement (SEM) and the smallest detectable change (SDC) (21,22,29). The SEM is derived by taking the square root of the mean square error term from the ANOVA when computing the ICC (22,29,30). The SEM is the estimate of the error associated with the patient's obtained score when compared with the hypothetical true score, and can be used to estimate a 95% confidence interval (CI) for the true score (30). The SEM was considered small if it represented less than 10% of the absolute scale range (31). The SDC was calculated using the formula $1.96 \times \sqrt{2} \times SEM$ (29). The SDC is an estimate of the amount of difference for which anything smaller cannot be reliable distinguished from random error in the measurement when evaluating outcome (29). Additionally, we constructed Bland–Altman plots for the rater pairs as well as for the two time points (32). In this way, we could examine the direction of the differences around the zero line (i.e. systematic bias) and whether the error of measurement is dependent on the magnitude of the mean score (i.e. heteroscedasticity) (33,34). As the data points in a Bland–Altman plot represent different numbers of observations, heteroscedasticity may be difficult to determine (34). Therefore we constructed bar charts of the differences (34) and correlation plots of the absolute differences and the means (33). In the case of no evidence of heteroscedasticity, limits of agreement (LOA) were calculated using the formula: $d \pm 1.96\ SD_{diff}$, where d is the mean differences and SD_{diff}, the standard deviation of the differences (32). Assuming that the differences are normally distributed, it is expected that 95% of the differences will be within the LOA (32). The distribution of the differences was visually assessed using histograms.

Sample size

A sample size of 102 patients was estimated to obtain ICC > 0.75 with a lower confidence limit greater than 0.60. A power of 80% and an alpha of 0.05 were used (35). For the assessment of the absolute reliability parameters a sample size of at least 50 patients is recommended (29).

Results

Demographics and functional performance

A sample size of 102 patients was assessed. Demographics and the functional performance profile are presented in Table I.

Table II. Inter-rater reliability of the MISA-DK scales.

	Relative reliability		Absolute reliability			Bland-Altman - Limits of agreement		
MISA-DK scales (score range)	$ICC_{1.1}$[a]	95% CI	SEM	SEM%[b]	SDC	d ± SD_{diff}	95% LOA	d% within LOA[c]
Positioning (4–12)	0.61	0.47–0.72	1.2	15%	3.3	0.1 ± 1.7	–3.3–3.5	92%
Self-feeding skills (7–21)	0.71	0.60–0.80	1.6	11%	4.4	0.4 ± 2.2	–4.0–4.8	96%
Liquid ingestion (7–21)	0.73	0.63–0.81	1.3	9%	3.6	–0.3 ± 1.9	–4.0–3.5	94%
Solid ingestion (12–36)	0.73	0.62–0.81	2.4	10%	6.7	–0.3 ± 3.4	–7.1–6.6	96%
Texture management–solids (8–24)	0.74	0.63–0.81	2.2	14%	6.1	–1.0 ± 2.9	–6.9–4.8	94%
Texture management–liquids (5–15)	0.76	0.66–0.83	1.4	14%	3.9	–0.2 ± 2.0	–4.1–3.7	94%
MISA total scale (43–129)	0.84	0.77–0.89	5.7	7%	15.8	–1.3 ± 8.0	–17.2–14.7	94%

Notes: [a]$p < .001$; [b]SEM as a percentage of the absolute scale range; [c]percentages of the differences within the LOA.

Relative reliability

The $ICC_{1.1}$ values for inter-rater reliability relating to the MISA-DK subscales and the MISA-DK total scale were in the range 0.61–0.84, indicating good to excellent inter-rater reliability (Table II). The $ICC_{3.1}$ values for intra-rater reliability relating to the MISA-DK subscales and the MISA-DK total scale were in the range 0.84–0.93, indicating excellent intra-rater reliability (Table III).

Absolute reliability

For inter-rater reliability, the SEM range was 1.2–5.7. The SEM represents 9–15% of the absolute scale range for the subscales and 7% of the absolute MISA-DK total scale range. The SDC range was 3.3–6.7 for the MISA-DK subscales and was 15.8 for the MISA-DK total scale (see Table II). For intra-rater reliability, the SEM range was 0.7–3.7. The SEM represents 6–10% of the absolute scale range for the subscales and 4% of the absolute MISA-DK total scale range. The SDC range was 1.9–4.4 for the MISA-DK subscales and was 10.3 for the MISA-DK total scale (see Table III).

The Bland–Altman plots for inter-rater reliability did not reveal any systematic bias. Indication of heteroscedasticity was present, but it could not be verified with bar charts or correlation plots. This is exemplified for the positioning scale in Figure 1.

The Bland-Altman plots for intra-rater reliability did not reveal any systematic bias or indication of heteroscedasticity. For inter-rater reliability, the expected 95% of the differences was within the LOA for two of the MISA-DK scales (see Table II). For intra-rater reliability, the expected 95% of the differences was within the LOA for 4 MISA-DK scales (see Table III).

Discussion

We evaluated the relative and absolute inter- and intra-rater reliability of the MISA-DK among a geriatric sample admitted to general medicine wards. In order to establish the measurement equivalence of the MISA-DK, we calculated the ICC to estimate relative reliability. We found excellent relative inter- and intra-rater reliability for the MISA-DK total scale. For the intra-rater reliability, we found excellent $ICC_{3.1}$ values (0.84–0.93 against 0.69–0.96) for all

Table III. Intra-rater reliability of the MISA-DK scales.

	Relative reliability		Absolute reliability			Bland-Altman – Limits of agreement		
MISA-DK scales (score range)	$ICC_{3.1}$[a]	95% CI	SEM	SEM%[b]	SDC	d ± SD_{diff}	95% LOA	d% within LOA[c]
Positioning (4–12)	0.87	0.83–0.90	0.7	9%	1.9	0.1 ± 1.0	–1.9–2.1	92%
Self-feeding skills (7–21)	0.86	0.82–0.89	1.1	8%	3.0	–0.1 ± 1.5	–3.1–2.9	92%
Liquid ingestion (7–21)	0.89	0.85–0.91	0.9	6%	2.5	0.0 ± 1.3	–2.6–2.6	95%
Solid ingestion (12–36)	0.88	0.84–0.91	1.6	7%	4.4	–0.3 ± 2.2	–4.7–4.1	98%
Texture management–solids (8–24)	0.84	0.80–0.88	1.6	10%	4.4	0.0 ± 2.3	–4.6–4.6	91%
Texture management–liquids (5–15)	0.88	0.85–0.91	1.0	10%	2.8	0.1 ± 1.4	–2.7–2.9	96%
MISA total scale (43–129)	0.93	0.90–0.94	3.7	4%	10.3	–0.4 ± 5.3	–11.0–10.2	95%

Notes: [a]$p < .001$; [b]SEM as a percentage of the absolute scale range; [c]percentages of the differences within the LOA.

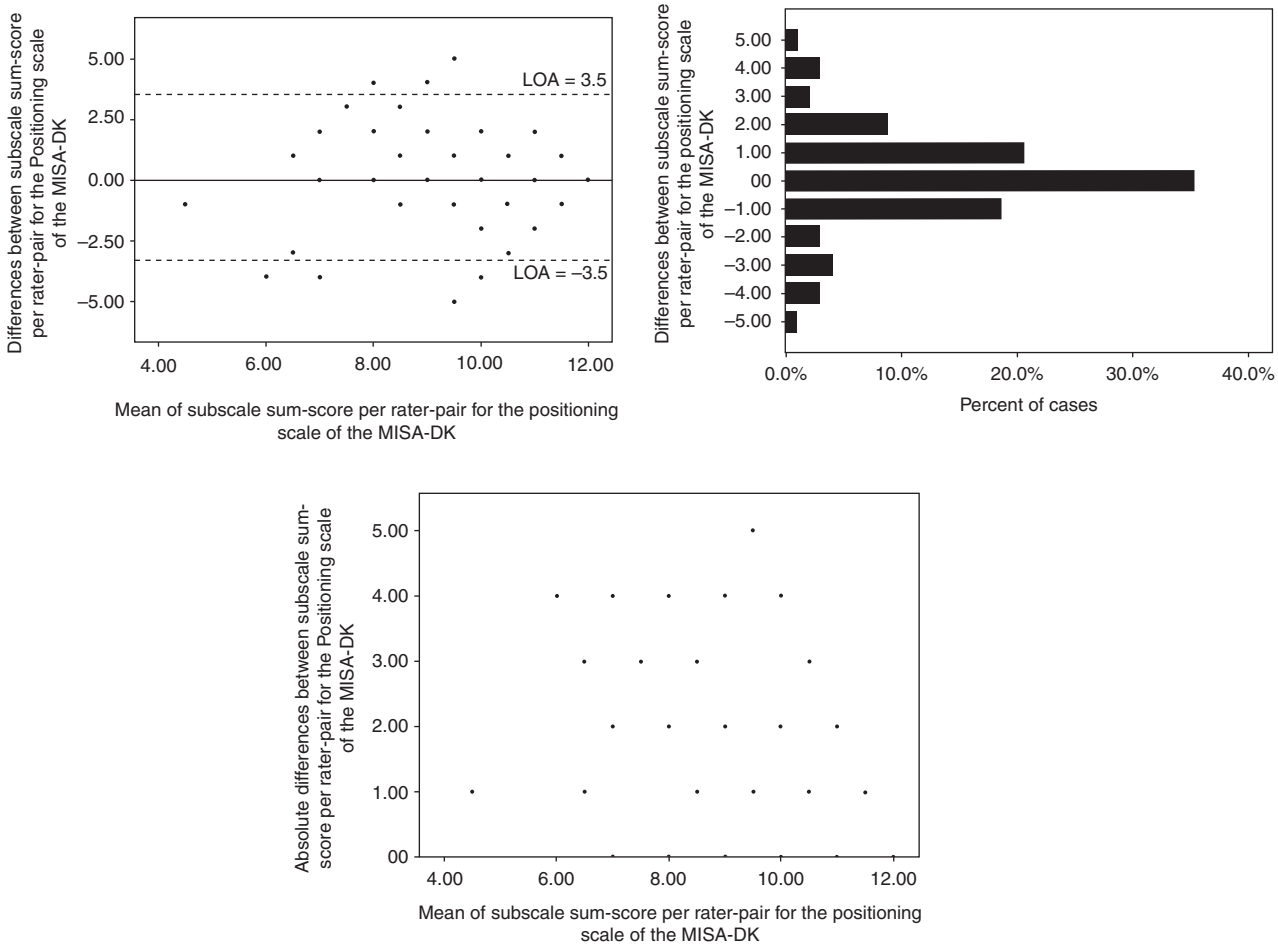


Figure 1. Bland–Altman plot, bar chart, and correlation plot of the differences between raters for the positioning subscale sum-score of the MISA-DK. A. In the Bland–Altman plot, each individual dot represents more than one data point. The sample distribution seems to be skewed to the right, indicating that the patients were rated to have a relative high ability level at the positioning scale. Heteroscedasticity appeared to be present as the magnitude of the differences seems to depend on the mean of subscale sum-score per rater. B. The bar chart revealed that for 86.2% of the cases, the absolute differences were less than ± 2 scale points of a total scale sum-score of 8 points. C. The correlation plot of the absolute differences and the means shows no association between the magnitude of the differences and the magnitude of the total scale sum-score.

scales, which equals the former Canadian validation study (14). However, for the MISA-DK subscales we found lower $ICC_{1.1}$ values for the inter-rater reliability (0.61–0.76 against 0.68–0.88). A possible explanation is that the ICC is strongly influenced by the variance of the trait in the sample (20,29,30). If the measurement scale is applied to a homogeneous population the between-subject variance is small, which results in a low ICC (29,30,33). From the Bland–Altman plots for all the MISA-DK scales, the sample distribution was skewed to the right, which may indicate homogeneity of the sample (see the example in Figure 1A). If the sample had been more heterogeneous, the between-subject variance would had been higher, resulting in larger reliability estimates (29,30,33). Therefore, ICCs measured in different populations might not be comparable (21). This implies that if the MISA-DK is to be administered among patients who differ from the sample in this study, then new reliability testing is required.

We extended the evaluation of the reproducibility of the MISA-DK and calculated the SEM, SDC, and LOA. This has not been addressed in earlier studies and our results add to the psychometric evidence of the MISA-DK. For the MISA-DK total scale, the estimated measurement error could be considered small as the SEM represents less than 10% of the absolute scale range in repeated measurement between raters as well as between time points by the same rater. However, at subscale level, the SEM exceeds 10% of the absolute scale ranges in repeated measurements between raters for four subscales. This may indicate less precision in repeated measurement by different raters than by the same rater. Using the SEM to estimate a 95% CI around an eight-point score on the positioning scale reveals

that the true score could range from 6 to 10 points in repeated measurement between raters and from 7 to 9 points in repeated measurement by the same rater. Additionally, the magnitude of the SDC was higher for the absolute inter-rater reliability than the absolute intra-rater reliability. This implies that in the case of using different raters, large score differences are required to exceed change and a true difference between measurements may be difficult to detect. This has to be considered if the MISA-DK is used for the purpose of evaluating outcome in research as well as in clinical practice.

We also constructed the Bland–Altman plots and calculated the LOA. For repeated measurement between rater-pairs, the LOA were wider than repeated measurement between time points by the same rater. In addition, more than 5% of the differences were outside the LOA for most of the MISA-DK scales in repeated measurement between rater-pairs. Post hoc analysis of the relative inter-rater reliability removing cases with differences outside the LOA revealed $ICC_{1.1}$ values in the range of 0.75 to 0.87 for the seven MISA-DK scales. This indicates excellent inter-rater reliability and is more similar to the Canadian validation study (14).

The inter-rater reliability was lower than the intra-rater reliability. This may be due to the fact that model 1 of the ICC provides a more conservative estimate of reliability than the other models (19) or because the ratings in our study may have been influenced by different sources of error. Three potential sources of errors may be present in any assessment, namely: the rating scales; the rating procedure; and the raters (36). For the rating scale, the trait may not be clearly defined or the rating scale categories may be ambiguously worded or insufficiently differentiated, which may result in inconsistent ratings (36). This could be a contributory explanation. In the former construct validation study of the MISA-DK, we found some problems with the scale categories when applying Rasch analysis (18). For the rating procedure, it could be argued that the judgment of the patients' performance based on video-recordings may be difficult compared with in-person judgment. However, if this was the explanation we would have expected to find poor intra-rater reliability as well, which was not the case. It seems more likely that rater effects have influenced the inter-rater reliability. This seems to be supported by the post hoc analysis of the relative inter-rater reliability discussed earlier. Although the raters received the same training, there seems to be some variability between raters. Plausible explanations could be differences in the interpretation of the operational scoring categories, in the degree of severity or leniency exhibited when scoring the patients' performance, and in the understanding and use of the ratings scale categories (36). These differences could be influenced by the raters' clinical experience and post-graduation in dysphagia management and this has to be investigated further. Investigation of rater effects and differential rater function can be realized using the many-facet Rasch measurement approach (36).

Methodological considerations

Our training course consisted of lectures, video-observation and scoring, facilitated discussion, and practice in the rater's own clinical context. This is in accordance with proposed learning approaches for rater training (37) with strategies based on experiential learning and reflection (38). Rater competence is operationalized in terms of conceptual knowledge and observation skills applied to a complex perceptual and cognitive measurement process (36). It may therefore be a difficult task to completely avoid rater error through training (39). Nevertheless, our training course might benefit from further development.

We recognize the inherent limitation of the use of video-recordings. The judgment of a clinician watching performance during a meal on a videotape is not an exact reflection of the judgment made in person in the clinical setting. We believe, however, that the advantages of video-recordings to ensure stability of the participants' performance across multiple testing and to ensure independence among raters exceed the disadvantages. Additionally, the large numbers of raters could have influenced our results and it would have been preferable to have fewer raters, but it was not realizable. On the other hand, in clinical practice we cannot be sure that the same limited sets of therapists provide services to the patients. So, in this light, our results may reflect the clinical reality in which the MISA-DK is to be implemented.

For the statistical analyses, we treated the sum-scores of the MISA-DK scales as continuous data. However, sum-scores based on ordinal scale levels are ordinal and not continuous (40). Differences of one point do not have the same meaning throughout the continuum when using ordinal scores (40–42). The use of parametric statistics with multi-item measurements has been the source of a longstanding debate (41). To overcome this dilemma, the ordinal rating scale data could be converted into equal interval measurements using the Rasch model (41,42). Therefore, it would be beneficial to apply the Rasch model to data obtained by the Danish and the Canadian version of the MISA. In this way, parametric statistics could be applied with confidence for further establishment of the measurement equivalence (41–43).

For the absolute reliability, we considered the SEM to be small if it represented less than 10% of the absolute scale range. However this criterion is arbitrary and other criteria may be used depending on the purpose of the measurement in question. Finally,

investigating heteroscedasticity using visual examination of plots could be questioned. Very slight heteroscedasticity could be overlooked, resulting in wider LOA for small differences than necessary and narrower LOA for large differences (32,33).

Conclusion

The relative reliability of the MISA-DK equals the original Canadian version with good to excellent inter-rater reliability and excellent intra-rater reliability. However, when extending the evaluation of the reproducibility of the MISA-DK, we found relatively large measurement errors and weaker precision when measurements are repeated by different raters than by the same rater. This has to be considered if the MISA-DK is to be used as outcome measure in research and in clinical practice. Further investigation of the rater effects on the MISA-DK scores as well as investigation of the measurement equivalence using the Rasch model is recommended.

Acknowledgements

The authors would like to thank the Research Council at Herlev University Hospital, the Research Foundation of the Danish Occupational Therapy Association (FF2/09-1) and the Lundbeck Foundation (FP03/2011) for financial support which made this research possible. The authors are grateful to all the participating patients, the facilities' staff, and all the occupational therapists who participated as raters. They would like to express special thanks to occupational therapist Charlotte Ehlers Hansen for assistance throughout in the data collection.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Cabre M, Serra-Prat M, Palomera E, Almirall J, Pallares R, Clave P. Prevalence and prognostic implications of dysphagia in elderly patients with pneumonia. Age Ageing 2010;39:39–45.
2. Rofes L, Arreola V, Romea M, Palomera E, Almirall J, Cabre M, et al. Pathophysiology of oropharyngeal dysphagia in the frail elderly. Neurogastroenterol Motil 2010;22:851–9.
3. Medin J, Larson J, Von Arbin M, Wredling R, Tham K. Elderly persons' experience and management of eating situations 6 month after stroke. Disabil Rehabil 2010;32:1346–53.
4. Perry L, McLaren S. Coping and adaptation at six months after stroke: Experiences with eating difficulties. Int J Nurs Stud 2003;40:185–95.
5. Perry L, McLaren S. An exploration of nutrition and eating disabilities in relation to quality of life at 6 months post-stroke. Health Soc Care Community 2004;12:288–97.
6. Treats TT. Use of the ICF in dysphagia management. Semin Speech Lang 2007;28:323–33.
7. Cichero JYA, Murdoch BE. Dysphagia: Foundation, theory and practice. Chichester: Wiley; 2006.
8. American Occupational Therapy Association. Specialized knowledge and skills in feeding, eating, and swallowing for occupational therapy practise. Am J Occup Ther 2007;61:686–99.
9. Hansen T, Kjærsgaard A, Faber J. Measuring elderly dysphagic patients' performance in eating: A review. Disabil Rehabil 2011;33:1931–40.
10. Lambert HC, Gisel EG, Wood-Dauphine S, Groher ME, Abrahamowicz M. McGill Ingestive Skills Assessment: Test manual and evaluation forms. Ottawa: Canadian Association of Occupational Therapists; 2006.
11. Lambert HC, Gisel EG, Groher ME, Wood- Dauphine SM. McGill Ingestive Skills Assessment (MISA): Development and first field test of an evaluation of functional ingestive skills of elderly persons. Dysphagia 2003;18:101–13.
12. Leopold NA, Kagel MA. Dysphagia – ingestion or deglutition? A proposed paradigm. Dysphagia 1997;12:202–6.
13. Canadian Association of Occupational therapy. CAOT Position Statement: Feeding, eating and swallowing and occupational therapy (2010). Available from http://www.caot.ca/default.asp?pageid=3948.
14. Lambert HC, Gisel EG, Groher ME, Abrahamowicz M, Wood-Dauphine S. Psychometric testing of the McGill Ingestive Skills Assessment. Am J Occup Ther 2006;60:409–19.
15. Lambert HC, Abrahamowicz M, Groher ME, Wood-Dauphine SM, Gisel EG. The McGill Ingestive Skills Assessment predicts time to death in an elderly population with neurogenic dysphagia: Preliminary evidence. Dysphagia 2005;20:123–32.
16. Hansen T, Lambert HC, Faber J. Content validation of a Danish version of "The McGill Ingestive Skills Assessment" for dysphagia management. Scand J Occup Ther 2011;18:282–93.
17. Streiner D, Norman G. Health measurement scales. 3rd ed. Oxford: Oxford University Press; 2003.
18. Hansen T, Lambert HC, Faber J. Validation of the Danish version of the McGill Ingestive Skills Assessment using classical test theory and the Rasch model. Disabil Rehabil. [Epub Oct 29 2011] Available from: http://informahealthcare.com/. doi:10.3109/09638288.2011.624249.
19. Portney IG, Watkins MP. Foundations of clinical research: Application to practice. 3rd ed. Upper Saddle River, NJ: Prentice Hall Health; 2009.
20. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. J Clin Epidemiol 2011;64:96–106.
21. Bruton A, Conway JH, Holgate ST. Reliability: What is it and how is it measured. Physiotherapy 2000;86:94–9.
22. De Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. J Clin Epidemiol 2006;59:1033–9.
23. Kjærsgaard A. Ansigt, mund og svælg: Undersøgelse og behandling efter Coombes konceptet [Face, mouth and throat: assessment and treatment according to the Coombes concept]. Copenhagen: FADL; 2005.
24. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index in stroke rehabilitation. J Clin Epidemiol 1989;42:703–9.
25. 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;12:189–98.

26. Bakke M, Bergendal B, McAllister A, Sjögreen L, Åsten P. Development and evaluation of a comprehensive screening for orofacial dysfunction. Swed Dent J 2007;31:75–84.
27. Smithard DG, O'Neill PA, Park C, England R, Renwick DS, Wyatt R, et al. Can bedside assessment reliably exclude aspiration following acute stroke? Age Ageing 1998;27:99–106.
28. Cicchetti DV, Bronen R, Spencer S, Haut S, Berg A, Oliver P, et al. Rating scales of measurement, issues of reliability: Resolving some critical issues for clinicians and researchers. J Nerv Ment Dis 2006;194:557–64.
29. Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. Am J Phys Med Rehabil 2005; 84:719–23.
30. Weir JP. Quantifying test–retest reliability using the Intraclass Correlation Coefficient and the SEM. J Strength Cond Res 2005;19:231–40.
31. Van Baalen B, Odding E, van Woensel MPC, van Kessel MA, Roebroeck ME, Stam HJ. Reliability and sensitivity to change of measurement instrument used in a traumatic brain injury population. Clin Rehabil 2006;20:686–700.
32. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
33. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217–38.
34. Smith MW, MA J, Stafford RS. Bar charts enhance Bland–Altman plots when value ranges are limited. J Clin Epidemiol 2010;63:180–4.
35. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med 1998;17:101–10.
36. Myford CM, Wolfe EW. Detecting and measuring rater effects using many-facet Rasch measurement: Part 1. J Appl Meas 2003;4:386–422.
37. Woehr DJ. Huffcutt AI Rater training for performance appraisal: A quantitative review. J Occup Organ Psychol 1994;67: 189–205.
38. Hounsgaard L, Eriksen JJ. Læring i sundhedsvæsenet [Learning within health care]. Copenhagen: Gyldendal; 2003.
39. Gingerich A, Regehr G, Eva KW. Rater-based assessments as social judgments: Rethinking the etiology of rater errors. Acad Med 2011;86(Suppl):S1–7.
40. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989;70:857–60.
41. Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. Rating scales as outcome measures for clinical trials in neurology: Problems, solutions, and recommendations. Lancet Neurol 2007;6: 1094–105.
42. Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Care Res 2007;57:1358–62.
43. Tennant A, Penta M, Tesio L, Grimby G, Thonnard JL, Slade A, et al. 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;42:1–37.