Development and validation of an algorithm to convert FIM® and EBI to an ADL score

2 March 2018

Charité – Universitätsmedizin Berlin
Institute of Medical Sociology and Rehabilitation Science

Email: anq-messplan@charite.de

Table of contents

Summary

Introduction: The aim of the National Rehabilitation Survey is to compare outcomes from all participating rehabilitation clinics. In order to register functional capacity in important areas of life in neurological and other rehabilitation programmes – and since 2016 in musculoskeletal rehabilitation programmes – either the FIM® instrument (FIM®) or the Extended Barthel Index (EBI) have been employed as outcome indicators. To make it possible to compare outcomes at all rehabilitation clinics irrespective of the instrument employed, FIM® or EBI scores need to be converted to a comprehensive Activities of Daily Living score (ADL score). Charité – Universitätsmedizin Berlin was commissioned by ANQ to undertake a study to develop and validate a novel ADL score.

Methodology: In the first phase of this mixed methods study, experts developed an algorithm whose function in converting FIM® and EBI scores to an ADL score was verified. In a second phase, the use of this conversion algorithm was validated in an observational study by determining FIM® *and* EBI scores for 263 patients attending four clinics or groups of clinics at five locations on admission to a neurological rehabilitation programme. In a third phase, the conversion algorithm was finalised and a consensus with regard to the algorithm was reached by the participating experts and researchers at the Charité based on the results of the validation study. The fourth and final phase involved a plausibility check of the ADL score using the data from the 2015 National Rehabilitation Survey for neurological rehabilitation.

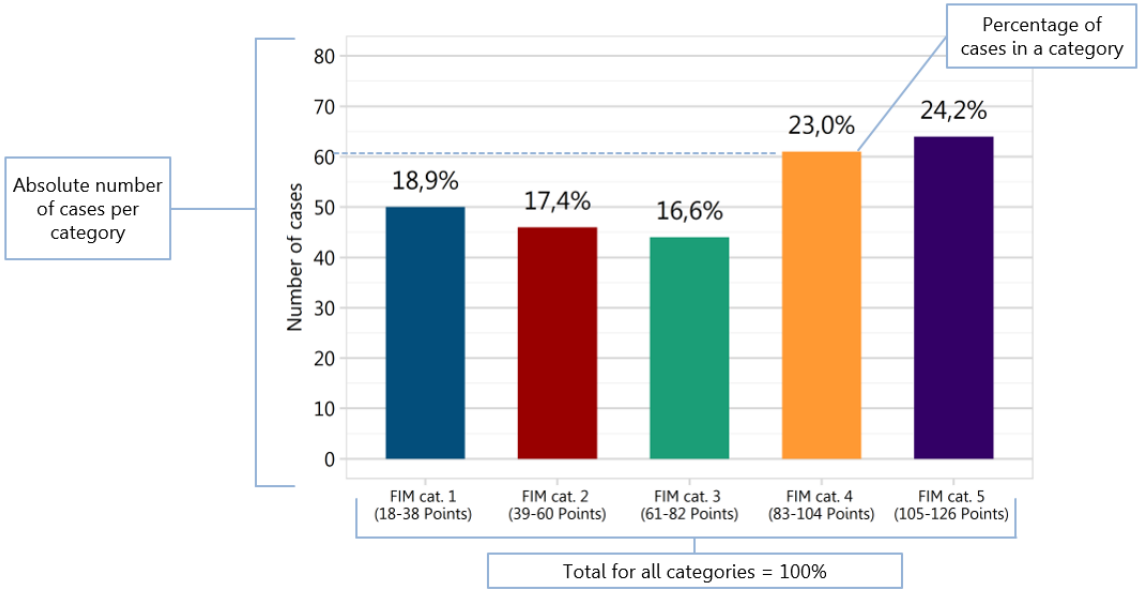
Results: Of the EBI items, 15 of the total of 16 were matched, usually in pairs, to the 18 FIM® items on the basis of their content. For each item pair, the 7 FIM® response categories were then matched to the 3 to 5 EBI response categories and scored on a scale of 0 to 4 points. For the ADL score, a scale of 0 to 60 points is used. The validation study revealed that the conversion algorithm was largely effective. Based on the results of the study, the conversion algorithm was modified slightly for some items and a consensus was reached at an expert workshop. Applying the conversion algorithm to the data obtained in the 2015 National Rehabilitation Survey demonstrated that the resultant ADL scores exhibited a high level of plausibility.

Discussion: There is no evidence that the novel ADL score reveals a substantial difference, either positive or negative, in the quality of clinics in terms of the measuring instrument employed. Regardless of whether FIM® or EBI are employed, a comparison of outcomes from all clinics within a specific therapeutic area now seems feasible and meaningful. The novel ADL score can be employed in the future as an indicator of outcomes of neurological, musculoskeletal and other rehabilitation programmes. At the same time, the empirically-obtained scores (FIM® or EBI) need also to be descriptively presented.

Interpretation aids for figures

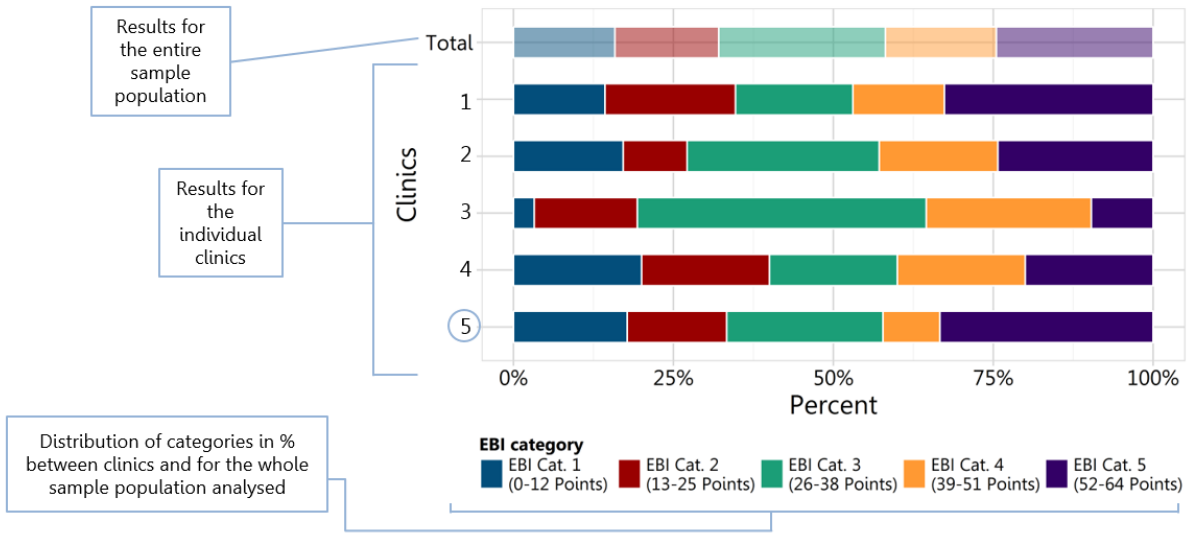
The following interpretation aids are intended to help readers understand the most important figure types employed in this report. For a list and explanation of technical terms, please refer to the glossary. Other types of figure are illustrated in the comparison reports of the National Rehabilitation Survey (e.g. Brünger et al., 2017).

Column chart (→ Glossary)



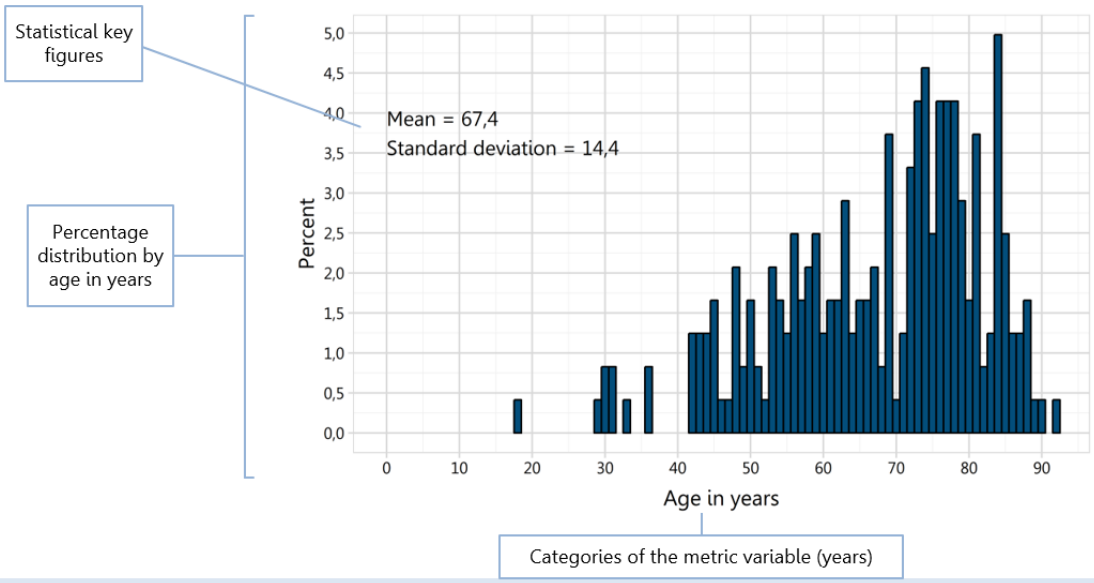
Example:
23% of cases (approx. 60 cases) have an FIM® score that is assigned to category 4 (83-104).

Stacked bar chart (→ Glossary)



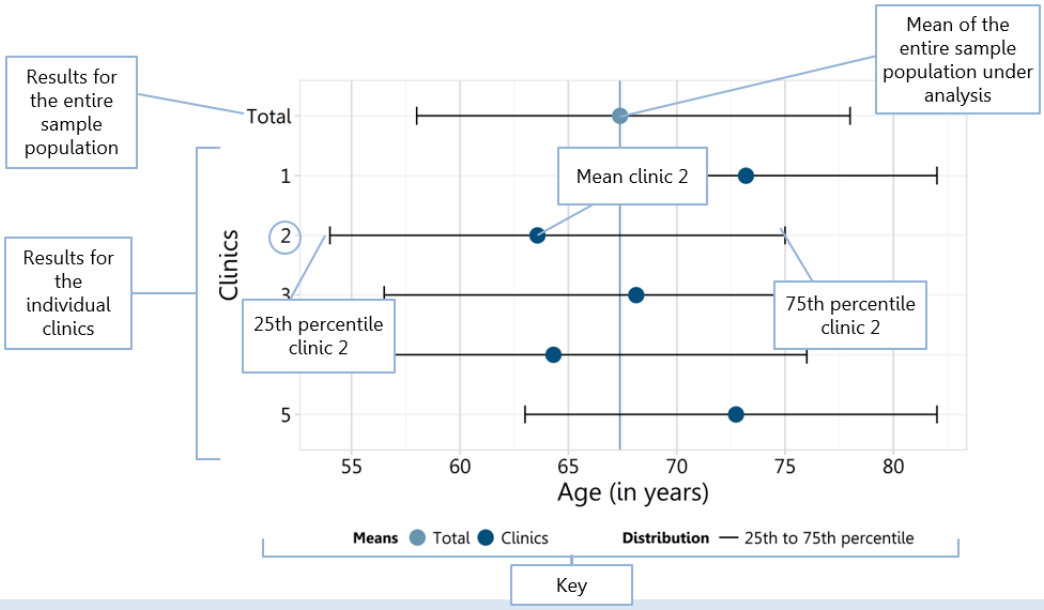
Example:
Approx. 20% of cases in clinic 5 were in EBI category 1, approx. 15% in category 2, approx. 25% in category 3, 10% in EBI category 4 and 30% in category 5. The proportion in category 4 was lower and in category 5 higher than in the entire study population.

Histogram (→ Glossary)



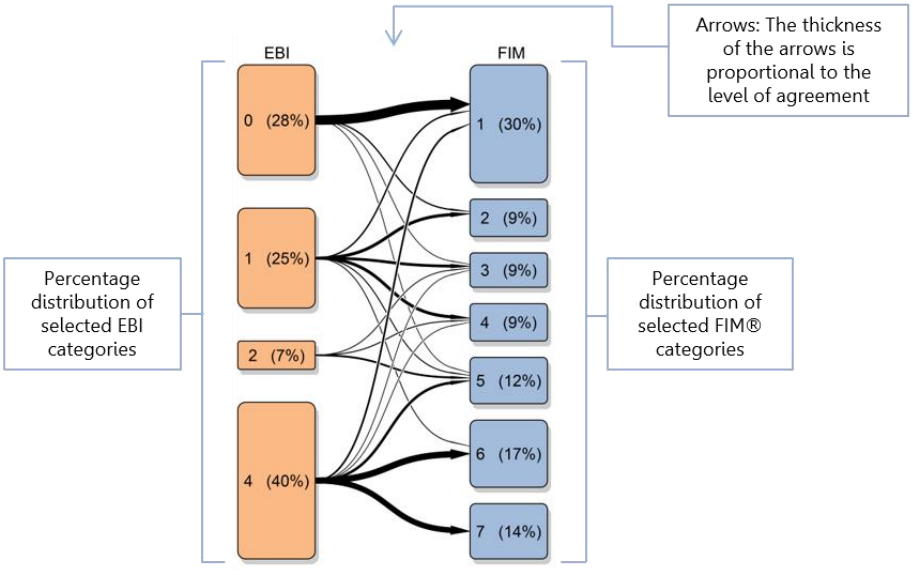
Example:
The mean (average) age is 68.2 years. The standard deviation is 11.2. Approx. 0.7% of patients were 44 years old (see annotation).

Simplified box plot (→ Glossary)



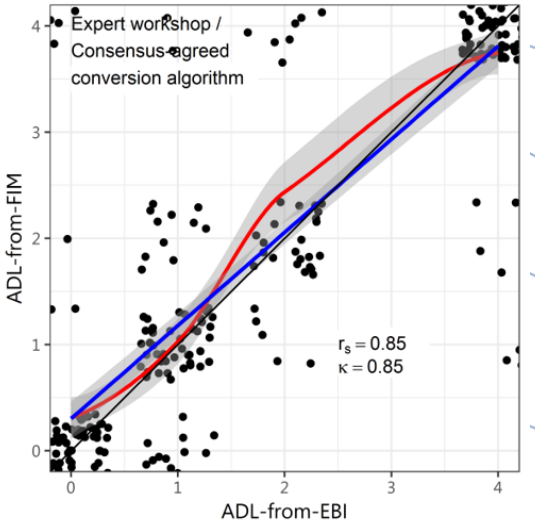
Example:
In clinic 43 the mean age was approx. 63 years. 25% of cases were a maximum of approx. 56 years old (25th percentile), 75% of cases were a maximum of approx. 71 years old (75th percentile). The total mean age was approx. 67.5 years old.

Transition plot (→ Glossary)



Example:
EBI category 4 was selected in 40% of cases. Most cases correspond with either an FIM® score of 6 or of 7. However there are some cases in which an EBI score of 4 corresponds to an FIM® score of 1, 2, 3, 4 or 5.

Scatter plot (→ Glossary)



ADL scores calculated from FIM®

Black diagonal line: hypothetical perfect agreement (correlation = 1))

Blue line: regression curve calculated by means of OLS regression

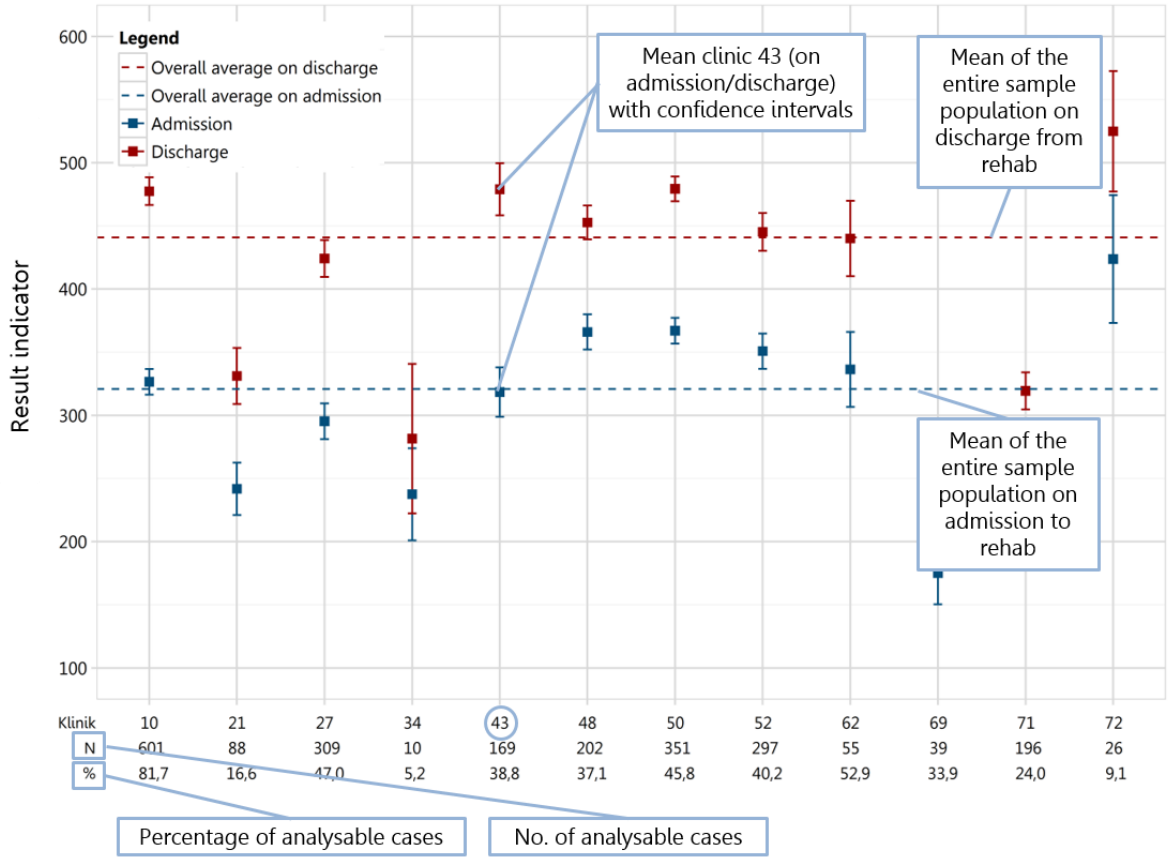
Red line: LOWESS regression curve

r_s: Spearman's rank correlation
K: weighted kappa

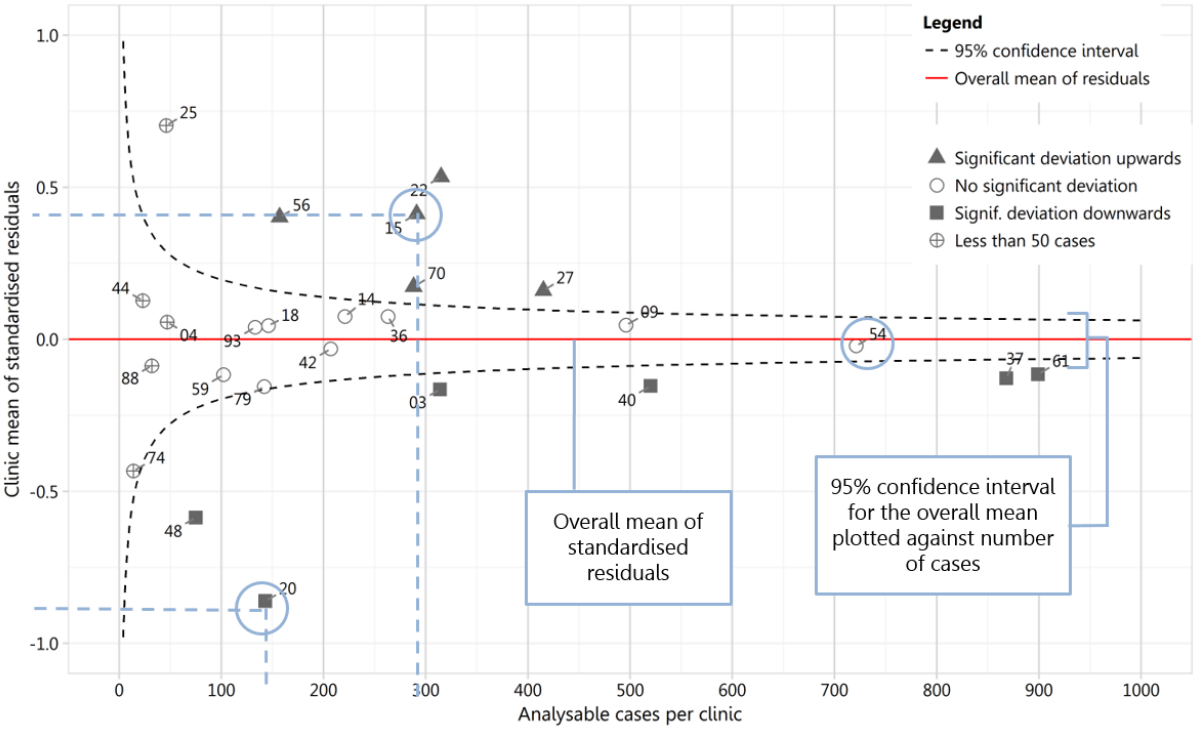
ADL scores calculated from EBI

Example:
The scores for ADL calculated from EBI and ADL calculated from FIM® are often identical. Nevertheless, ADL-from-EBI score 1 is, noticeably, often selected at the same time as ADL-from-FIM® scores 2 or 3. The OLS regression curve and the LOWESS regression curve indicate a high level of agreement due to their relative closeness to the black line.

Error bar chart: Mean scores with 95% confidence intervals (→ Glossary)



Example:
In clinic 43 the mean outcome indicator on admission was approx. 320 points and at discharge approx. 490 points. With a certainty of 95%, the true mean score on admission was in a range of 300 to 330 (confidence interval). As the confidence intervals on admission and discharge do not overlap, the discharge score is significantly higher than the admission score. 169 cases were included in the analysis for clinic 43. The percentage of cases that could be scored of the total cases registered by clinic 43 was 38.8%.
The average total scores on admission and discharge are indicated by dashed lines.



Example:
The mean standard residual in the case of clinic 27 was 0.41. In view of the number of cases (n≈290) and the individual case mix, the result for this clinic was significantly better than the expected result.
The mean standard residual for clinic 20 was -0.88. In view of the number of cases (n≈140) and the individual case mix, the result from this clinic was significantly worse than the expected result.
The mean standard residual score for clinic 54 was within the confidence interval of the overall mean score and is therefore not statistically significant different from the overall mean.

1. Introduction

The Swiss National Association for Quality Development in Clinics and Hospitals (ANQ) has been undertaking quality surveys of inpatient rehabilitation throughout Switzerland since 2013 (Menzi, 2015). The National Rehabilitation Survey has been designed in the form of a full survey. The survey assesses the success of rehabilitation in the modules "neurological rehabilitation" and "other rehabilitation programmes", documenting the purposes of participation and the targets to be achieved, while employing either the FIM® instrument (FIM®) or the Extended Barthel Index (EBI) to record the quality of outcomes. Since 2016, FIM® and EBI have also been employed in connection with musculoskeletal rehabilitation, replacing the Health Assessment Questionnaire (HAQ) (Menzi *et al.*, 2017).

In the measurement year 2015, 25 clinics employed FIM® and 6 clinics EBI in connection with the evaluation of neurological rehabilitation (Brünger *et al.*, 2017). In 2015, FIM® was employed in 41 clinics, and EBI in 6 clinics in other rehabilitation programmes (Schlumbohm *et al.*, 2017). In 2016, 37 clinics employed FIM® and 20 clinics EBI in connection with musculoskeletal rehabilitation (data not yet published).

Both instruments measure functionality in important Activities of Daily Living (ADL) (Wade, 1992). FIM® was developed from 1983 onward (Granger *et al.*, 1986) and comprises 18 items relevant to differing activities of life, each with 7 identically formulated levels of function, from "Complete independence" (7 points) to "Total assistance" (1 point). The range of total possible scores is thus 18 to 126 points (Keith *et al.*, 1987).

EBI represents an extended version of the Barthel Index originally developed in 1965 (Collin *et al.*, 1988; Heuschmann *et al.*, 2005) as it adds six items for communication and/or cognition (Prosiegel *et al.*, 1996). In the case of EBI, the evaluation categories for each item were also further differentiated. EBI comprises 16 items each with 3 to 5 response categories, which are formulated differently for each item. Each item is assessed on a scale of 0 (greatest level of impairment) to 4 points (lowest level of impairment), resulting in a total score of between 0 and 64 points. Both instruments also sub-categorise items as either motoric or cognitive (Prosiegel *et al.*, 1996). Quality of outcomes is compared on the basis of the global overall score (Menzi *et al.*, 2017).

Both instruments also have in common that scoring is undertaken in practice by therapeutic or nursing staff who observe patients during normal daily clinical routine. If required, this can be supplemented by interviewing patients or relatives. Completing an EBI evaluation takes somewhat less time than in the case of FIM® (Prosiegel et al., 1996). FIM® and EBI are structurally very similar, exhibit close correlation and are comparable in terms of their reliability (Prosiegel *et al.*, 1996; Schädler *et al.*, 2006). However, EBI is focussed on determining the level of dependence on external help while FIM® is instead more closely aimed at recording patient functional independence (Schädler *et al.*, 2006).

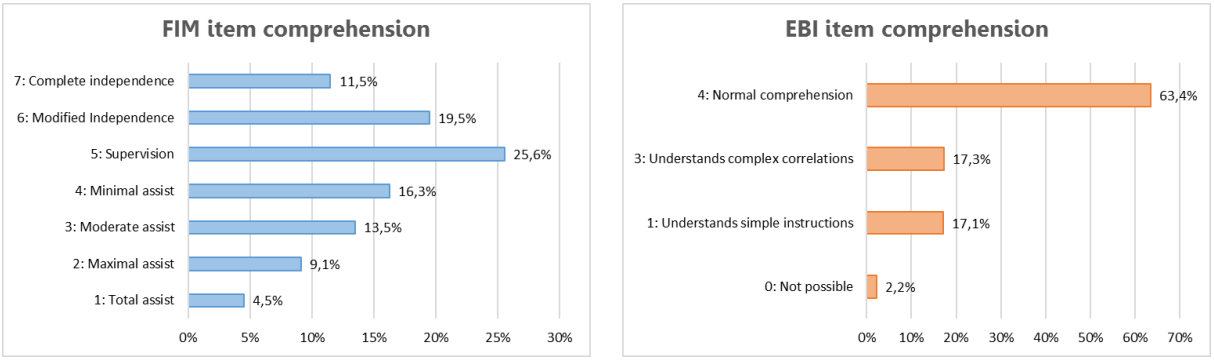
The aim of the National Rehabilitation Survey is to enable outcomes from all clinics dealing with a given therapeutic area to be compared, irrespective of whether they employ FIM® or EBI. This requires FIM® and EBI scores to be converted to a common scaled score.

Preliminary studies were carried out at the Charité to analyse the comparability of FIM® and EBI. This was complemented by a literature search for further approaches to establishing ADL scores on the basis

of FIM® and EBI scores. The results of these analyses were presented at meetings of the ANQ Rehabilitation Quality Committee on 19 March 2015 and on 11 June 2015.

FIM® and EBI not only have a differing number of response categories, the content-related design of the response categories also varies. Ceiling effects tend to occur more frequently in EBI than FIM® (see example item "Comprehension", Figure 1).

Figure 1: Distribution of response categories for the item "Comprehension" (Data from ANQ: Neurological Rehabilitation 2014)



Consideration of the operationalisation of individual items with regard to their content suggests that low categories are selected less frequently in EBI than in FIM® due to the broadness of definition of its intermediate categories. This can be seen in the distribution of individual EBI and FIM® items (see e.g. Figure 1).

As part of the preliminary analyses, FIM® and EBI overall scores were normed on a scale of 0 to 100 points, in order to evaluate the comparability of the two instruments. This revealed the mean EBI score was considerably lower than that of FIM®. These results are consistent with the findings of the KIQ (predecessor of the ANQ) pilot project.

All preliminary analyses were undertaken using data from the National Rehabilitation Survey, in which either FIM® or EBI scores were documented. No data are available for cases in which both instruments were simultaneously employed to assess the same rehabilitation patients. In order to compare outcomes at clinics irrespective of the instrument used, a validated algorithm that will convert FIM® and EBI values to a common ADL score is required.

A literature search revealed no specific evidence of an existing, scientifically validated form of ADL score that would effectively represent a common overall score for both EBI and FIM®. There have only been attempts to date to formulate a common motoric score by combining the Barthel index and the FIM® motoric sub-category (Nyein et al., 1999; Kwon et al., 2004). In addition, there are studies demonstrating a high level of agreement between and/or convertibility of FIM® and Barthel Index (Gosman-Hedstrom, Svensson, 2000; Turner Stokes *et al.*, 2010; Prodinger *et al.*, 2017) or that compare the psychometric properties of both instruments (Hsueh *et al.*, 2002; Vanbellingen *et al.*, 2016). However, these also do not take into account the socio-cognitive items. Another study demonstrates that FIM® and the Barthel Index can be assigned to common ICF Core Sets (Grill *et al.*, 2006).

There is, therefore, currently no scientifically valid ADL scoring system that can evaluate outcomes from clinics employing either FIM® or EBI. For this reason, at the present time there is no feasible and meaningful method to compare outcomes at all rehabilitation clinics dealing with a particular therapeutic area independently of the assessment method employed. To date, clinics using FIM® or EBI have been assessed separately with regard to quality of outcomes (Brünger *et al.*, 2017). A fair comparison of outcomes in clinics irrespective of whether they use FIM® or EBI will require the development and empirical validation of a novel algorithm that will make it possible to convert FIM® and EBI scores to a common ADL score. At its meeting of 11 June 2015, the Rehabilitation Quality Committee decided to implement a study to develop a suitable algorithm. The Institute of Medical Sociology and Rehabilitation Science at the Charité – Universitätsmedizin Berlin was subsequently commissioned by ANQ to undertake such a study.

The study's goal is to develop and verify a valid mapping algorithm that will generate an ADL score that will fairly compare outcomes at clinics independently of whether FIM® or EBI are employed.

2. Methodology

2.1. Study design

The study was divided into several phases:

Firstly, within the framework of an expert workshop a method was developed to convert FIM® and EBI with regard to their content to a common ADL score. *Secondly*, individual rehabilitation clinics scored patients using both FIM® and to EBI so that the conversion algorithm thus generated with regard to content could be empirically validated. *Thirdly*, the algorithm used to convert FIM® and EBI to an ADL score based on the results of the study was finalised and a consensus was agreed upon by the participating experts. *Fourthly*, a plausibility check of the ADL-conversion algorithm was performed using data from the 2015 National Rehabilitation Survey.

This mixed methods approach, employing qualitative and quantitative methods, makes it possible to develop and test a validated conversion algorithm, justified in terms of its content, on the basis of both expert opinion and empirical evidence.

The study was performed in accordance with the ethical principles for medical research stipulated in the Declaration of Helsinki. In order to participate in the study, patients and experts were required to provide their written informed consent. Study documentation for patients was provided in German, French and Italian and, for experts, in German and French. The Ethics Committee of the Canton of Bern (KEK) was approached in order to establish its competence with regard to study approval. The KEK stated that it was not accountable with regard to study approval. The Data Protection Officers and the Ethics Committee of the Charité - Universitätsmedizin Berlin approved the study (EA1/218/16). The study was registered with the international study register of the U.S. National Institutes of Health, ClinicalTrials.gov (NCT03233789).

For reasons of anonymity, the naming of individual persons or rehabilitation clinics was avoided in this report. The IDs used for clinics participating in the study differ from those used in the national comparison reports.

2.2. Development of the ADL score

On 7 April 2016 an all-day expert workshop was held in Bern to formulate a content-related method to convert FIM® and EBI results to a novel ADL score. The participants were seven experts who had practical experience of the use of FIM® and/or EBI or had appropriate expertise with regard to these instruments (Table 1). Relevant professional groups and different institutions from various language regions were involved. The workshop was interpreted simultaneously into and from German and French. The Charité prepared the programme for the expert workshop, and also moderated it and undertook the necessary follow-up. To facilitate evaluation of the expert workshop, the discussion was recorded with a digital tape recorder.

The aim of the workshop was to examine if and which FIM® and EBI items were compatible in terms of content and which response categories of the FIM® and EBI item pairs corresponded in terms of content. On this basis, a first algorithm, justified in terms of content, was developed to convert FIM® and EBI scores to an ADL score the quality of which could be determined in the subsequent validation study.

Table 1: Participants in the expert workshops

Name	Institution	Profession/post
Mathieu Courcelle	Institution de Lavigny	Head of Department of Ergotherapy, FIM® trainer
Claudia Gabriel	Lucerne Cantonal Hospital	Nursing care expert
Myrta Kohler	Valens Clinics	Registered nurse, Head of Nursing Development; Institute of Applied Nursing Sciences at the St. Gallen University of Applied Sciences
Julien Moncharmont	Lausanne University Hospital	Head of Department of Ergotherapy, FIM® trainer
Anke Steel-Sailer	Swiss Paraplegic Centre, Nottwil	Senior Consultant for Paraplegiology, Head of Research Rehabilitation Quality Management
Claude Vaney	Bern Montana Clinic	Senior Consultant for Neurology
Bernadette Vögele	Lucerne Cantonal Hospital	Physiotherapist, Head of Ambulant Therapy Neurorehabilitation

2.3. Validation of ADL scores

2.3.1. Study design and data collection

An observational study was carried out with the aim of validating, and, if required, modifying on the basis of empirical evidence the algorithm to convert FIM®/EBI values to an ADL score developed at the expert workshop. To achieve this, patients undergoing inpatient neurological rehabilitation were evaluated using both the FIM® and EBI methods. It was decided that evaluations were to be undertaken on admission to rehabilitation as the range of variation of functional ability is broader at this point in time than at discharge from rehabilitation and fewer ceiling effects occur (Brünger *et al.*, 2017). The National Rehabilitation Survey requires that neurological rehabilitation patients must be evaluated using at least one of the two instruments (FIM® or EBI) – in other words, only one non-routine instrument was additionally required for the study.

Patients were recruited consecutively for the study who were undergoing inpatient neurological rehabilitation, were aged 18 years or older, were admitted into one of the study clinics at the beginning of January 2016 and who gave their consent to participate in the study either in person or through an appointed legal representative. The final study data was collected on 24 February 2017 (Table 2). Inadequate language skills (German, French or Italian) were defined as an exclusion criterion in terms of the provision of informed consent. As FIM® and EBI scores were both recorded by observers, no added time or effort was expected of the study patients.

To make the results of the study more widely applicable, a multicentre approach was chosen. Thus, four clinics/groups of clinics were recruited at a total of five sites (Table 2). To determine whether there were any effects on evaluation processes due to prior experience with using either FIM® or EBI, among the four study clinics recruited were some that routinely employed FIM® and one clinic that was using EBI in the National Rehabilitation Survey during the study period.

The aim was to include at least n=198 complete cases/data in the study. This figure was determined on the basis of the inter-rater reliability coefficient kappa. This requires a total of at least $2K^2$ cases, whereby *K* represents the number of levels of function specified for each item (Sim, Wright, 2005). FIM® items have 7 response categories, while EBI (and therefore the novel ADL to be developed) has up to 5 possible levels of function. The minimum numbers of cases required are thus: FIM®: $n=2 \times 7^2=98$; both EBI and ADL: $n=2 \times 5^2=50$. In order to preclude the possibility of an insufficient number of values (including data from the National Rehabilitation Survey), a figure of 70 cases per clinic (group) was deemed necessary, giving a total of n=280 cases. The number of patients finally recruited was n=263, the distribution of which among clinics can be seen in Table 2.

Table 2: Participating rehabilitation clinics, dates of data collection and number of cases

Clinic/group of clinics	Site	FIM®/EBI*	Data collection start	Data collection finish	Number of cases
Clinica Hildebrand Centro di riabilitazione	Brissago TI	FIM®	18 Jan. 2016	25 Dec. 2016	70
Klinik Bethesda	Tschugg BE	FIM®	16 Feb. 2016	24 Feb. 2017	43
Valens Clinics	Valens SG	EBI	27 Jan. 2016	4 Aug. 2016	70
RehaClinic	Bad Zurzach AG & Kilchberg ZH	FIM® FIM®	8 Jan. 2016 8 Jan. 2016	20 Jun. 2016 20 Apr. 2016	31 49

* Employed routinely during the study period for the National Rehabilitation Survey

Clinics with personnel that could record functional capacity using both instruments were specifically selected for the study. Due to their complexity, the EBI and FIM® instruments require experience and/or training in their use. This complexity makes it particularly important that data for both EBI and FIM® are collected strictly in accordance with guidelines to ensure their validity (Prosiegel *et al.*, 1996; Uniform Data System for Medical Rehabilitation, 2009). All the neurological rehabilitation clinics already employ either FIM® *or* EBI in the framework of the National Rehabilitation Survey and are extremely experienced in the use of the instrument they routinely employ. To ensure that all study clinics were able to collect valid data using both FIM® and EBI, ANQ offered clinics training for the study in the use of the instrument that they were asked to additionally employ.

To exclude inter-rater effects, FIM® and EBI scores for each patient were recorded by identical personnel/teams of personnel. To reduce the dependence of the scores taken on the judgement of individual

personnel, at least two different persons in each clinic/group of clinics were involved in evaluating patients using both FIM® and EBI. Between 3 and 5 persons recorded scores for the study in each clinic/group of clinics, a total of 16 individuals.

In order to generate a valid conversion algorithm for all levels of impairment, those rehabilitation patients sampled included patients at all levels of impairment. Pre-analysis showed that, in practice and depending on the rehabilitation clinic, severely impaired patients were likely to be underrepresented in individual clinics (see also Brünger *et al.*, 2017). Recruitment of rehabilitation patients at each clinic was therefore stratified according to impairment level. Five impairment levels for FIM® and EBI were therefore categorised, each covering a similar range of impairment (FIM®: 21-22 points, EBI: 13 points). In each participating clinic the aim was to recruit n=14 rehabilitation patients from each impairment category. Assignment to impairment categories was made on the basis of scores on admission to rehabilitation. In clinics routinely employing FIM®, these scores were recorded using FIM®, and for EBI clinics, using EBI (Table 3).

Table 3: Impairment categories in FIM®/EBI and the target number of cases per clinic

Category (Impairment)	FIM® points 18 - 126	EBI points 0 - 64	Recruitment n per clinic
1 (high)	18 - 38	0 - 12	14
2	39 - 60	13 - 25	14
3	61 - 82	26 - 38	14
4	83 - 104	39 - 51	14
5 (low)	105 - 126	52 - 64	14

The study data were sent regularly by the study clinics to the Charité at intervals of approximately two to three months. This made it possible to check the completeness of the data at the time it was being collected, perform plausibility analyses and provide study clinics with feedback on the data being delivered. The study clinics were informed if data was identified that seemed implausible. In such cases, the data were corrected to ensure FIM® and EBI were correctly documented, as per the procedural handbook (Menzi *et al.*, 2017) and the manuals for FIM® (Uniform Data System for Medical Rehabilitation, 2009) and for EBI (Prosiegel et al., 1996).

The study data were linked with the routinely collected data from the National Rehabilitation Survey by means of a case identification number (CID). The survey data were sent to the Charité biannually at set dates (28 February and 31 August) (Charité - Universitätsmedizin Berlin, 2017). This included the minimum data required by the Swiss Federal Statistical Office (SFSO) in the form of socio-demographic information such as gender, age, nationality and insurance status; and information on rehabilitation such as primary and secondary diagnoses and duration of rehabilitation treatment (Bundesamt für Statistik,

2016). In addition, any co-morbidity was also recorded on admission to rehabilitation using the Cumulative Illness Rating Scale (CIRS) (Linn *et al.*, 1968). Depending on the clinic, FIM® or EBI scores were also recorded on discharge from rehabilitation. Detailed information on the data collected in the National Rehabilitation Survey is included in the procedural handbook (Menzi *et al.*, 2017).

2.3.2. Data analysis

After all data had been sent and checked for plausibility, an analysis was performed at the level of individual items and/or item pairs and at the level of the ADL score – calculated on the basis of FIM® and/or EBI, in order to validate the mapping algorithm developed at the expert workshop. Beforehand, a description of the sample population was prepared.

The responses for the FIM® and EBI item pairs were represented in graph form as "transition plots". Each FIM® response category (a total of 7 possible categories) and each EBI response category (a total of 3–5 categories per item) is represented by a column. The length of the column indicates the frequency of the particular response category in the sample population. The FIM® and EBI columns are linked by arrows, indicating which EBI category/ies were selected in the presence of a given FIM® category. The thickness of the arrows represents the frequency of each combination of a FIM® and an EBI response category. The more frequently an EBI category was selected in combination with a particular FIM® category, the thicker the arrow.

A second type of graph employed takes the form of scatter plots, which in this case show the ADL score calculated on the basis of FIM® score on the y-axis and the ADL score calculated on the basis of EDI score on the x-axis. Each point represents a patient's two ADL scores (calculated on the basis of FIM® and EBI scores). Ideally, all the points would lie on the straight black line from 0 to 4 (ADL score range at item level) on both axes, with a gradient of 1. The actual regression curve is shown in blue. The additional red regression curve was generated using the LOWESS (locally weighted scatter plot smoothing) method (Cleveland, 1979). This represents an alternative way of modelling the data, which – in simple terms – calculates an individual slope for each point to obtain the regression curve. The LOWESS curve can, therefore, deviate from a straight line and is instead curved with a varying gradient along its course. If the ADL scores from FIM® and from EBI were to perfectly agree, this red line would also be a straight line from 0 to 4 with a gradient of 1.

In addition to the representation of the results in graph form, levels of agreement with ADL scores (calculated on the basis of FIM® and EBI scores) were determined for each mapped pair of corresponding FIM® and EBI items. These included the percentage agreement of the ADL scores calculated on the basis of FIM® and EBI scores, the level of agreement per the kappa (κ) coefficient, the correlation per Spearman's rank correlation coefficient (r_s) and the difference in points between the mean ADL scores calculated on the basis of FIM® and EBI scores.

The percentage agreement can be in the range 0% to 100%. The higher the value, the better the agreement. The results for Cohen's kappa are in a range 0 to 1. Values for kappa of 0.41–0.60 can be interpreted as moderate agreement, while values of 0.61–0.80 indicate close agreement, and values of 0.81–

1.00 represent (almost) perfect agreement (Landis, Koch, 1977). This report details weighted kappa values (Cohen, 1968). To take into account items on an ordinal scale, the square of the difference between both assessed variables is employed. The result for Spearman's rank correlation is always in the range -1 to +1; in this case, 0 indicates no linear correlation, -1 indicates perfect negative correlation and +1 indicates perfect positive linear correlation between the ADL scores and the underlying FIM® scores and ADL scores and underlying EBI scores.

At item level in this study, ADL scores obtained on the basis of FIM® and EBI scores can differ by -4 to +4 points (ADL score level -60 to +60); in the ideal case, the difference would be 0 points.

If various ways of mapping response categories were proposed for individual item pairs at the expert workshop, such values were calculated, and scatter plots were generated for each variant (labelled "Expert workshop *variant 1*", "Expert workshop *variant 2*" etc.). If the data indicated that other content-related mappings of response categories seemed to fit better (See Results section 3.2), levels of agreement were also reported for the alternative conversion algorithms (labelled "validation study").

At the scale level, a scatter plot with regression curves was also generated. The closer the regression curve to the ideal black line from 0 to 60 (ADL score range) with a gradient of 1, the closer is the agreement between the ADL score and the underlying FIM® score and the ADL score and the underlying EBI score. For purposes of comparison, the conversion algorithm from the first expert workshop and the algorithm agreed upon at the consensus workshop are shown.

Figures for the three variants are given in the tables showing the percentage agreement, the weighted kappa value, the intraclass correlation coefficient (ICC), Spearman's rank correlation coefficient, the gradient of the regression curves and the adjusted coefficient of determination R^2. The ICC is, in the same way as the kappa coefficient, a measure of reliability, with a value between -1 and +1. It is particularly well suited to assessing the level of agreement between (quasi-)interval-scaled instruments, in this case the agreement between ADL scores and the underlying FIM® and EBI scores. ICC was calculated using the "absolute agreement" method. This represents the more stringent assessment variable. The coefficient of determination R^2 can be interpreted as the quality of the prediction of the reliability of the ADL score calculated on the basis of the FIM® score or EBI score (and vice versa); the resultant value is in the range 0 to 1. The higher the value of R^2, the better the quality of the prediction. Adjusted R^2 takes into account the number of independent variables.

In addition, the difference in the mean values of the ADL scores calculated on the basis of the FIM® and EBI scores was calculated. This difference was statistically analysed by parameters using the t-test for dependent samples.

2.4. Expert consensus on the ADL score

The results of the validation study were presented at a second half-day expert workshop in Bern on 28 March 2017. With one exception due to timetabling difficulties, the same experts were present (compare Table 1). Some of the mapping proposals made by the experts were revised based on the empirical data and the Charité presented these new proposals for discussion. All item pairs and response category

mappings were discussed, some of the classifications slightly modified in view of the empirical results and due to content-related considerations, and a final consensus agreement was reached on the conversion algorithm for each of the 15 item pairs.

All analyses were then performed once again for the final conversion algorithm agreed by consensus using the same procedure described in section 2.3.2 (labelled "Consensus-agreed conversion algorithm").

2.5. Plausibility check of the ADL score

Finally, the algorithm agreed by consensus was applied to the neurological data of the 2015 National Rehabilitation Survey.

To begin with, the non-adjusted mean FIM®, EBI and ADL scores for each clinic on admission to and discharge from rehabilitation were plotted on an error bar chart with confidence intervals of 95%. The dashed lines show the mean value for all clinics on admission to and discharge from rehabilitation.

In addition, a risk-adjusted evaluation was performed. This was done by generating three separate linear regressions respectively for FIM®, EBI and ADL scores at discharge from rehabilitation as the dependent/outcome variable. The independent/predictor variables, taking into account the patient structure at the individual clinics, included the confounders and the scores for the outcome variable at admission (Table 4). This methodological approach to compare outcomes is also cited in the national comparison report (Brünger *et al.*, 2017).

Table 4: Confounders and data sources

Confounders	Data sources
Gender Age Nationality Principal diagnosis per ICD-10 (discharge) Duration of treatment Insurance status Principal cost carrier Place of residence before admission Place of residence after discharge	SFSO statistics: minimum data
Baseline status: level of impairment on admission (t1)	FIM®, EBI or ADL score on admission calculated from FIM® or EBI
Co-morbidity	CIRS: Cumulative Illness Rating Scale

As in the national comparison reports, funnel plots are employed as visual tools. For each clinic, the mean of the standardised residuals is plotted against the number of cases employed in the analysis for

each clinic. In this way, all possible relationships between outcomes and clinic sizes are presented. Clinics with significantly better outcomes appear above the upper confidence interval of the total mean value (represented by grey triangles). Clinics whose scores were significantly lower than the risk-adjusted expected values appear below the lower confidence interval (represented by grey squares). Clinics with outcomes that would be expected in view of the case mix are represented by plain circles. Clinics with less than 50 scorable cases are represented by crossed circles (Brünger *et al.*, 2017).

The descriptive results, both in the form of error bar charts and funnel plots, allow scores to be compared: on the one hand ADL scores, and on the other, FIM® and EBI scores. It must be stated that the comparability is limited as the basis for comparison is different in each case. Comparing quality of outcome on the basis of ADL scores makes possible a comparison of a clinic with all the other clinics involved in the particular therapeutic area. Comparing outcomes on the basis of the FIM® scores, on the other hand, only makes possible a comparison of a clinic with other clinics employing FIM®. This is also the case for EBI. Assuming the clinics using EBI or FIM® are systematically different in terms of their patient structure and also in terms of changes to outcome indicators from admission to discharge, this can lead to different results for individual clinics, depending on whether ADL scores or FIM®/EBI overall scores are employed at discharge. Divergent results are more likely for clinics using the instrument that is used in fewer cases in a given therapeutic area (for all three areas – neurological rehabilitation, musculoskeletal rehabilitation, and other rehabilitation programmes – this is currently EBI).

It should be noted that the analyses of data from the National Rehabilitation Survey in this report were only performed in order to verify the plausibility of the ADL conversion algorithm. The final outcome quality comparisons for 2016 are expected to be published in the national comparison reports at the end of 2018.

3. Results

3.1. Development of an ADL score: expert workshop

It was first discussed which EBI items corresponded in terms of content with which FIM® items. This led to a consensus that the mapping as depicted in Table 5 might be a feasible option. One point of note is that three EBI items are each mapped to two FIM® items. EBI item 3 (Dressing) is mapped to FIM® items D (Dressing upper body) and E (Dressing lower body). EBI item 4 (Bathing/showering/taking a complete sponge bath) were mapped to FIM® items C (Bathing) and K (Transfer to shower/tub). In addition, EBI item 8 (Getting on and off toilet) and FIM® items F (Toileting) and J (Transfer to toilet) correspond. No FIM® item could be mapped to EBI item 16 (Spatial neglect). In contrast, all FIM® items could be mapped to an EBI item, yielding a total of 15 FIM®-EBI item pairs.

Table 5: Mapping the individual EBI and FIM® items.

EBI: 16 Items - Scoring 4-0		FIM: 18 Items - Scoring 1-7		
1.	Eating/drinking	A	Eating	
2.	Grooming	B	Grooming	
3.	Dressing/undressing	D E	Dressing - Upper Body Dressing - Lower Body	!
4.	Bathing/showering/taking a complete sponge bath	C K	Bathing Transfer: Tub/Shower	!
5.	Transfer: wheelchair/bed & vv.	I	Transfer: Bed/Chair/Wheelchair	
6.	Locomotion on level surfaces	L	Locomotion: Walk/Wheelchair + additional item	
7.	Ascending/descending stairs	M	Locomotion: Stairs	
8.	Use of the toilet	F J	Toileting Transfer: Toilet	!
9.	Bowel management	H	Bowel Management	
10.	Bladder Management	G	Bladder Management	
11.	Cognitive comprehension	N	Comprehension + additional item	
12.	Expression	O	Expression + additional item	
13.	Social Interaction	P	Social Interaction	
14.	Problem Solving	Q	Problem Solving	
15.	Memory/learning/orientation	R	Memory	
16.	Seeing - Neglect	-	-	

The second task was to discuss for each item pair which FIM® and EBI response categories were congruent in terms of content and could thus be transferred to a common ADL item. Several variants per item pair were sometimes developed by the experts. As FIM®, with 7 possible response categories, is more differentiated than EBI, with 3 to 5 response categories, FIM® response categories were mapped to corresponding EBI response categories (and not vice versa). At the item level, ADL is scored on a range from 0 (lowest functional capacity) to 4 points (greatest functional capacity). Like EBI, up to 5 levels of function are specified for each ADL item (possible scores 0, 1, 2, 3 and 4 – whereby only the levels 0, 2 and 4 exist for all items). On the basis of 15 FIM®-EBI item pairs, this means that, in total, ADL has a score range of 0 to 60 points.

Detailed below are the conversion algorithms proposed by the experts, sometimes with several variants, for each item pair.

Item 1: Eating and drinking

For item pair 1, "Eating and drinking", the experts proposed two variants for how the FIM® and EBI response categories could be converted into a novel ADL score. As shown in Table 6, there are only four response categories in EBI (category 1 does not exist), which need to be mapped to 7 categories in FIM®. As is standard practice in the case of FIM®, a score in categories 1 to 5 indicates how much assistance a patient requires from another person, while a score in categories 6 and 7 indicates how well a patient can perform an activity without a helper.

Table 6: ADL item 1 Eating and drinking: Comparison of EBI item 1 and FIM® item A

EBI: 1. Eating and drinking		FIM: A. Eating		
Not possible. Or: Enteral nutrition (PEG or nasal tube) that the patient cannot operate independently.	0	1	Total assist (Patient = less than 25%)	Helper
		2	Maximal assist (Patient = 25% or more)	
Food needs to be prepared (e.g. meat and vegetables need to be cut up).	2	3	Moderate assist (Patient = 50% or more)	
		4	Minimal assist (Patient = 75% or more)	
Patient can independently eat (unprepared) food with the help of aids e.g. breakfast board, thicker cutlery handles, etc. Or: Patient can operate gastric tube independently.	3	5	Supervision (Patient = 100%)	
		6	Modified Independence (Device)	No helper
Independent.	4	7	Complete Independence (timely, safely)	

Variant 1. As shown in Table 7, the experts proposed for this variant that the FIM® response categories "Total assistance" and "Maximal assistance" when eating and drinking should be considered equivalent to EBI category "Unable". In addition, the FIM® categories "Moderate assistance", "Minimal contact assistance" and "Supervision/setup" were seen as equivalent to the EBI category "Food must be prepared". In this variant, the EBI category "Feeding alone possible (without preparation) using auxiliary aids" is considered equivalent to the FIM® category "Modified independence". EBI "Independent" and FIM® "Complete independence" represent the category of greatest capacity.

Table 7: Conversion algorithm proposed by the experts – *variant 1*.

Expert workshop variant 1		
EBI 1	**ADL**	**FIM® A**
0	0	1 2
2	2	3 4 5
3	3	6
4	4	7

Variant 2. As shown in Table 8, in this variant the lowest EBI category ("Unable") corresponds to the lowest four FIM® categories ("Total assistance" though to "Minimum assistance"). The fifth category in FIM® ("Supervision/setup") corresponds in this case to EBI category 2, "Food must be prepared". EBI category 3 "Feeding alone possible using auxiliary aids" in this variant has no corresponding category in FIM®; in other words, a score of 3 can only be selected in ADL if this is based on an EBI score. "Modified independence" and "Complete independence" in FIM® are equivalent to the EBI category of greatest capacity "Independent".

Table 8: Item 1: Conversion algorithm proposed by the experts – *variant 2*.

Expert workshop variant 2		
EBI 1	**ADL**	**FIM® A**
0	0	1 2 3 4
2	2	5
3	3	
4	4	6, 7

Item 2: Personal hygiene

For item pair 2, "Personal hygiene", the experts proposed three variants for how the FIM® and EBI response categories could be converted to a novel ADL score. As shown in Table 9, for this item pair there are 7 response categories in FIM® that need to be mapped to four response categories in EBI.

Table 9: ADL item 2 Personal hygiene: Comparison of EBI item 2 and FIM® item B

EBI: 2. Grooming (washing face, combing, shaving, brushing teeth)	
Not possible.	0
Requires the help of an assistant with some, but not all, of these procedures.	1
Possible with a little help (e.g. removal of the cap from the toothpaste tube). Or: Does not require direct assistance, but needs prompting/encouragement/supervision with regard to certain procedures.	2
Independent personal hygiene is possible with aids, e.g. longer handle for comb, flannel, brush.	3
Independent (in all of the preceding areas); patients who, for instance, are unable to plait their hair true to style are to be nonetheless classed as independent.	4

FIM: B. Grooming (Oral care, washing hair, hands, face)		
1	Total assist (Patient = less than 25%)	Helper
2	Maximal assist (Patient = 25% or more)	
3	Moderate assist (Patient = 50% or more)	
4	Minimal assist (Patient = 75% or more)	
5	Supervision (Patient = 100%)	
6	Modified Independence (Device)	No helper
7	Complete Independence (timely, safely)	

Variant 1: As Table 10 shows, in this proposed variant the lowest EBI category ("Unable") corresponds to the lowest FIM® category "Total assistance (patient can manage less than 25% and requires assistance by a helper)". The EBI response "Needs assistance by a helper with some but not all activities" is considered to correspond to FIM® categories "Maximal assistance" and "Moderate assistance". "Needs some help" is mapped to FIM® categories "Minimal contact assistance" and "Supervision/setup". "Personal hygiene alone using auxiliary aids possible" is considered to correspond to the FIM® category "Modified independence" and "Independent" to the FIM® category "Complete independence".

Table 10: Item 2: Conversion algorithm proposed by the experts – *variant 1*

Expert workshop variant 1		
EBI 2	**ADL**	**FIM® B**
0	0	1
1	1	2 3
2	2	4 5
3	3	6
4	4	7

Variant 2: As shown in Table 11, the lowest EBI category "Unable" in this variant is mapped to the two lowest FIM® categories "Total assistance" and "Maximal assistance". The EBI response "Needs assistance by a helper with some but not with all activities" is considered to correspond to FIM® categories "Moderate assistance" and "Minimal contact assistance". "Needs some help" is considered to correspond to the FIM® category "Supervision/setup". As in variant 1, the EBI category "Personal hygiene alone possible using auxiliary aids" is mapped to the FIM® category "Modified independence", while the EBI category "Independent" corresponds to the FIM® category "Complete independence".

Table 11: Item 2: Conversion algorithm proposed by the experts – *variant 2*

Expert workshop variant 2		
EBI 2	**ADL**	**FIM® B**
0	0	1 2
1	1	3 4
2	2	5
3	3	6
4	4	7

Variant 3: As shown in Table 12, the lowest EBI category "Unable" is mapped to the lowest FIM® category "Total assistance". The EBI response "Needs assistance by a helper with some but not with all activities" is considered to correspond to the FIM® category "Maximal assistance". The EBI category "Needs some help" is mapped to FIM® categories "Moderate assistance", "Minimal contact assistance" and "Supervision/setup". In this variant there is no ADL category 3, and therefore the EBI category "Personal hygiene alone possible using auxiliary aids" and "Independent" correspond to FIM® categories "Modified independence" and "Complete independence".

Table 12: Item 2: Conversion algorithm proposed by the experts – *variant 3*

Expert workshop variant 3		
EBI 2	**ADL**	**FIM® B**
0	0	1
1	1	2
2	2	3 4 5
3 4	4	6 7

Item 3: Dressing and undressing

As seen in Table 13, in the third ADL item the seven FIM® response categories are mapped to the four possible responses in EBI (category 3 does not exist). In addition, mapping from FIM® involves two items: D. "Dressing upper body" and E. "Dressing lower body". In total, the experts proposed four variants. On the one hand, they proposed two differing mappings of the categories. On the other, they recommended that the two FIM® items be taken into account either by employing the mean score of both items or the lower of the two scores.

Table 13: ADL item 3 – Dressing and undressing: Comparison of EBI item 3 and FIM® items D and E

Dressing/undressing (including tying shoelaces and doing up buttons)	
Not possible.	0
Needs help with putting on or taking off most clothing items, but not all of them. Or: Participates effectively, but still needs help with putting on or taking off every item of clothing.	1
Needs help with just a few procedures (e.g.: help with tying shoelaces, doing up or undoing buttons, putting on elastic stockings or needs aids such as braces). Or: Does not require direct assistance, but needs prompting/encouragement/supervision with regard to certain procedures.	2
Independent. (use of stocking pullers, for instance, is allowed).	4

FIM: D. Dressing - Upper Body	E. Dressing - Lower Body		
1	1	Total assist (Patient = less than 25%)	Helper
2	2	Maximal assist (Patient = 25% or more)	
3	3	Moderate assist (Patient = 50% or more)	
4	4	Minimal assist (Patient = 75% or more)	
5	5	Supervision (Patient = 100%)	
6	6	Modified Independence (Device)	No helper
7	7	Complete Independence (timely, safely)	

Variant 1: As shown in Table 14, in this variant the lowest EBI category "Unable" is mapped to the lowest FIM® category "Total assistance". The EBI response "Needs help when dressing and undressing for most but not all clothing" is considered to correspond to FIM® categories "Maximal assistance" and "Moderate assistance". The EBI category "Needs help with only with a few procedures" is mapped to FIM® categories "Minimal contact assistance" and "Supervision/setup". The best EBI category "Independent" is considered to correspond to FIM® categories "Modified" and "Complete independence". The experts also proposed two subvariant mappings of the FIM® categories:

Variant 1A: The lower of the two values for item D and item E should be employed.

Variant 1B: The mean score should be employed, rounded down if the value is not a whole number.

Table 14: Item 3: Conversion algorithm proposed by the experts – *variant 1*

Expert workshop variant 1		
EBI 3	**ADL**	**FIM® D+E***
0	0	1
1	1	2
		3
2	2	4
		5
4	4	6
		7

* Variant 1A: lower of the scores for D and E
* Variant 1B: mean score of D and E (rounded down, if not a whole number)

Variant 2: Table 15 shows that in this variant the lowest EBI category "Unable" should correspond to an ADL score, but does not, however, correspond to an FIM® response category. Both of the lowest FIM® categories, "Total assistance" and "Maximal assistance", are considered to correspond to EBI category "Needs help when dressing and undressing for most but not all clothing". The EBI category "Needs help with only a few procedures" is mapped to FIM® categories "Moderate assistance", "Minimal contact assistance" and "Supervision/setup". The best EBI category "Independent" is considered to correspond to the FIM® categories "Modified independence" and "Complete independence". As in the case of variant 1, the experts proposed two subvariant mappings of the FIM® categories:

Variant 2A: The lower of the two values for item D and item E should be employed.

Variant 2B: The mean score should be employed, rounded down if the value is not a whole number.

Table 15: Item 3: Conversion algorithm proposed by the experts – *variant 2*

Expert workshop variant 2		
EBI 3	**ADL**	**FIM® D+E***
0	0	
1	1	1,2
2	2	3 4 5
4	4	6 7

* Variant 2A: lower of the scores for D and E
* Variant 2B: mean score of D and E (rounded down if not a whole number)

Item 4: Bathing/showering/taking a complete sponge bath

The experts proposed two variants for mapping the item "Bathing/showering/taking a complete sponge bath". As FIM® is mapped from two items (items C and K), employing *either* the lowest value for mapping *or* the rounded-down mean score was proposed. The seven response categories in FIM® need to be mapped to five EBI response categories (Table 16).

Table 16: ADL item 4 – Bathing/showering/taking a complete sponge bath: Comparison of EBI item 4 and FIM® items C and K

EBI: 4. Bathing/showering/taking a complete sponge bath	
Not possible.	0
Needs the help of an assistant for some, but not all, procedures (e.g. requires help with the necessary transfers or with towelling down; can wash the upper body but needs help with the lower parts of the body).	1
Possible with a little help (e.g. removal of caps from shampoo and shower gel bottles). Or: Does not require direct assistance, but needs prompting/encouragement/supervision with regard to certain procedures.	2
Use of aids is necessary (e.g. lift, bath or shower chair), although the patient can use them independently.	3
Independent.	4

FIM: C. Bathing			
	K. Transfers: Tub, Shower		
1	1	Total assist (Patient = less than 25%)	Helper
2	2	Maximal assist (Patient = 25% or more)	
3	3	Moderate assist (Patient = 50% or more)	
4	4	Minimal assist (Patient = 75% or more)	
5	5	Supervision (Patient = 100%)	
6	6	Modified Independence (Device)	No helper
7	7	Complete Independence (timely, safely)	

Variant 1: As shown in Table 17, in this variant the lowest EBI category "Unable" is mapped to the lowest two FIM® categories "Total assistance" and "Maximal assistance". The EBI response "Needs assistance by a helper with some but not with all activities" is considered to correspond to the FIM® category "Moderate assistance". In this variant, the EBI category "Needs some help" is mapped to FIM® categories "Minimal contact assistance" and "Supervision/setup". The EBI category "Auxiliary aid required" is considered to correspond to the FIM® category "Modified independence (using auxiliary aids)", while the EBI response category "Independent" corresponds to the FIM® response category "Complete independence". The experts also proposed two subvariant mappings of the FIM® categories.

Variant 1A: The lower of the two values for item C and item K should be employed.

Variant 1B: The mean score of FIM® items C and K should be employed, rounded down if the value is not a whole number.

Table 17: Item 4: Conversion algorithm proposed by the experts – *variant 1*

Expert workshop variant 1		
EBI 4	**ADL**	**FIM® C+K***
0	0	1 2
1	1	3
2	2	4 5
3	3	6
4	4	7

* Variant 1A: lower of the scores for C and K
* Variant 1B: mean score of C and K (rounded down if not a whole number)

Variant 2: As shown in Table 18, in this variant the lowest EBI category "Unable" is to be mapped to the lowest two FIM® categories "Total assistance" and "Maximal assistance". The EBI response categories "Needs assistance by a helper with some but not with all activities" is considered to correspond to FIM® categories "Moderate assistance" and "Minimal contact assistance". In this variant, the EBI category "Needs some help" is mapped to FIM® category "Supervision/setup". The EBI category "Auxiliary aid required" is considered to correspond to the FIM® category "Modified independence (using auxiliary aids)", while the EBI response category "Independent" corresponds to the FIM® response category "Complete independence". The experts also proposed two subvariant mappings of the FIM® categories:

Variant 2A: The lower of the two values for item C and item K should be employed.

Variant 2B: The mean score of FIM® items C and K should be employed, rounded down if the value is not a whole number.

Table 18: Item 4: Conversion algorithm proposed by the experts – *variant 2*

Expert workshop variant 2		
EBI 4	**ADL**	**FIM® C+K***
0	0	1 2
1	1	3 4
2	2	5
3	3	6
4	4	7

* Variant 2A: lower of the scores for C and K
* Variant 2B: mean score of C and K (rounded down if not a whole number)

Item 5: Bed to chair/wheelchair transfer

For item pair 5, "Bed to chair/wheelchair transfer" seven FIM® categories need to be mapped to four EBI categories, as EBI category 3 does not exist (Table 19).

At the expert workshop, after a fair amount of discussion the mapping given in Table 20 was proposed: the EBI category "Unable" should correspond to the FIM® categories "Total assistance" and "Maximal assistance". It was also suggested that the FIM® responses "Moderate assistance" and "Minimal contact assistance" should correspond to "Needs assistance by a helper with some but not all activities". The FIM® category "Supervision/setup" corresponds to the EBI category "No direct help required, but reminders/prompts/supervision for some activities". The highest EBI response category, "Independent", corresponds to the upper two FIM® response categories "Modified independence" and "Complete independence".

Table 19: ADL item 5 – Bed to chair/wheelchair transfer: Comparison of EBI item 5 and FIM® item I

EBI: 5. Transfer: wheelchair/bed and vice versa	
Not possible.	0
Requires the help of an assistant with some, but not all, of these procedures.	1
Does not require direct assistance, but needs prompting/encouragement/ supervision with regard to certain procedures (e.g. must be prompted to apply the brake).	2
Independent.	4

FIM: I. Transfer Bed/ Chair/Wheelchair		
1	Total assist (Patient = less than 25%)	Helper
2	Maximal assist (Patient = 25% or more)	
3	Moderate assist (Patient = 50% or more)	
4	Minimal assist (Patient = 75% or more)	
5	Supervision (Patient = 100%)	
6	Modified Independence (Device)	No helper
7	Complete Independence (timely, safely)	

Table 20: Item 5: Conversion algorithm proposed by the experts

Expert workshop		
EBI 5	**ADL**	**FIM® I**
0	0	1 2
1	1	3 4
2	2	5
4	4	6 7

Item 6: Locomotion

For item pair 6 "Locomotion", the experts proposed one variant for how the FIM® and EBI response categories be converted to the novel ADL score. As detailed in Table 21, the five response categories in EBI are to be mapped to seven FIM® categories. As is standard practice in the case of FIM®, a score in categories 1 to 5 indicates how much assistance a patient requires from another person, while a score in categories 6 and 7 indicates how well a patient can perform an activity without a helper. Furthermore, for this item an additional item needs to be determined: whether rehabilitation patients are mobile using only a wheelchair or both walk and use a wheelchair.

Table 21: ADL item 6 – Locomotion: Comparison of EBI item 6 and FIM® item L

Additional Item	
1	Walking
2	Wheelchair
3	Both

EBI: 6. Locomotion on level surfaces	
Not possible (cannot walk or use a wheelchair independently).	0
Requires a wheelchair or rollator, but can use it largely independently (i.e.: can cover longer distances, does not bump into obstacles, can navigate curves, turn around, etc. and only needs help in very few situations). Or: Patient can walk short distances (less than 50 m), but only with assistance.	1
Can manage shorter distances (less than 50 m) without assistance or the use of a handrail, but needs a wheelchair, rollator or supervision for longer distances (more than 50 m).	2
Can manage longer distances (more than 50 m) without a rollator or handrail, but needs a walking stick, crutch, brace or similar.	3
Can walk independently over longer distances without any assistance or aid.	4

FIM: L. Locomotion: Walk/Wheelchair		
1	Total assist (Patient = less than 25%)	Helper
2	Maximal assist (Patient = 25% or more)	
3	Moderate assist (Patient = 50% or more)	
4	Minimal assist (Patient = 75% or more)	
5	Supervision (Patient = 100%)	
6	Modified Independence (Device)	No helper
7	Complete Independence (timely, safely)	

As displayed in Table 22, the experts proposed two mapping variants, depending on whether the patient walks or uses a wheelchair. If the additional item "Both" is selected, the patient is scored as if the response was "Walks". Patients in a wheelchair can receive a maximum ADL score of 2, as can be seen in the right-hand section of the table, while patients who can walk can score up to 4 points.

The lower two ADL categories, 0 and 1, are allocated independently of the additional item: in the lowest ADL category, the EBI category "Unable" corresponds to the FIM® categories "Total assistance" and "Maximal assistance", while in the next highest EBI category, "Requires a wheelchair or rollator, which can be operated independently for the most part", corresponds to the FIM categories "Moderate assistance" and "Minimal contact assistance". If "Walks" is selected for the additional item, the EBI category "Can walk independently for short distances (<50 m) without helper or handrails" corresponds to the FIM category "Supervision". Furthermore, the experts proposed that the EBI category "Can walk independently for short distances (<50 m) without helper or handrails" correspond to FIM® category "Modified independence". Finally, the best EBI category "Can walk independently, even for long distances" is to correspond to the FIM category "Complete independence".

If "Wheelchair" is selected for the additional item, the highest ADL category is 2. This corresponds to the EBI category "Can walk independently for short distances (<50 m) and, in FIM®, the categories "Supervision/setup" and "Modified independence". EBI categories 3 and 4 and FIM® category 7 have been completely deleted in the case of this variant.

Table 22: Item 6: Conversion algorithm proposed by the experts for "Walks" and "Wheelchair"

Expert workshop variant Walking		
EBI 6	**Gehen**	**FIM® L**
0	0	1
		2
1	1	3
		4
2	2	5
3	3	6
4	4	7

Expert workshop variant Wheelchair		
EBI 6	**ADL**	**FIM® L**
0	0	1
		2
1	1	3
		4
2	2	5
		6

Item 7: Stairs

For item pair 7, "Stairs", the experts proposed one variant for how the FIM® and EBI response categories be converted to the novel ADL score. As detailed in Table 23, the four response categories in EBI are to be mapped to seven FIM® categories. Response category 3 in EBI does not exist. As usual in FIM®, categories 1 to 5 concern how much assistance a patient requires from another person, while a score in categories 6 and 7 indicates a patient can perform an activity without a helper.

Table 23: ADL item 7 – Stairs: Comparison of EBI item 7 and FIM® item M

EBI: 7. Ascending/descending stairs	
Not possible.	0
Possible, but only with the extensive assistance of another person (e.g. with lifting a leg).	1
Possible, but needs a little help or supervision from another person.	2
Can navigate stairs independently (holding on to a banister rail or using a walking stick, crutch or similar is permissible).	4

FIM: M. Locomotion: Stairs		
1	Total assist (Patient = less than 25%)	Helper
2	Maximal assist (Patient = 25% or more)	
3	Moderate assist (Patient = 50% or more)	
4	Minimal assist (Patient = 75% or more)	
5	Supervision (Patient = 100%)	
6	Modified Independence (Device)	No helper
7	Complete Independence (timely, safely)	

In the experts' proposal (see Table 24), the lowest EBI category "Unable" corresponds to the lowest FIM® category "Total assistance". If "Possible, but only with significant assistance by a helper (e.g. help in lifting a leg)" is selected in the EBI item "Moving up and down stairs", this should correspond to FIM® "Maximal assistance" and "Moderate assistance". It was also proposed that the EBI category "Possible, but only with some assistance or supervision by a helper" be mapped to the FIM® categories "Minimal contact assistance" and "Supervision". For the EBI item "Moving up and down stairs", the optimal category "Possible independently" is to correspond to the FIM® response categories "Complete independence" and "Modified independence".

Table 24: Item 7: Conversion algorithm proposed by the experts

Expert workshop		
EBI 7	**ADL**	**FIM® M**
0	0	1
1	1	2 3
2	2	4 5
4	4	6 7

Item 8: Toilet use

For item pair 8, "Toilet use", the experts proposed one variant for how the FIM® and EBI response categories be converted into the novel ADL score. As detailed in Table 25, the four response categories in EBI are to be mapped to seven FIM® categories (EBI category 3 does not exist). As is standard practice in the case of FIM®, a score in categories 1 to 5 indicates how much assistance a patient requires from another person, while a score in categories 6 and 7 indicates how well a patient can perform an activity without a helper. The FIM® score is determined by two FIM® items: item F "Toileting" and item J "Transfer to toilet".

Table 25: ADL item 8 – Toilet use: Comparison of EBI item 8 and FIM® items F and J

EBI: 8. Use of the toilet (transfer, dressing/undressing, personal hygiene, flushing)	
Not possible (cannot walk or use a wheelchair independently).	0
Needs the assistance of another person for certain, but not all, procedures (e.g. independent transfer, but needs help with dressing/undressing).	1
Does not require direct assistance, but needs prompting/encouragement/supervision with regard to certain procedures (e.g. needs prompting to flush the toilet).	2
Either independent or does not require assistance with these activities (e.g. has diapers or suprapubic cystostomy/in-dwelling catheter, therefore does not need to use the toilet).	4

FIM: F. Toileting	J. Transfer: Toilet		
1	1	Total assist (Patient = less than 25%)	Helper
2	2	Maximal assist (Patient = 25% or more)	
3	3	Moderate assist (Patient = 50% or more)	
4	4	Minimal assist (Patient = 75% or more)	
5	5	Supervision (Patient = 100%)	
6	6	Modified Independence (Device)	No helper
7	7	Complete Independence (timely, safely)	

Table 26 details how the experts proposed EBI and FIM® items are to be converted to a common ADL score. If "Toilet use" in EBI has been categorised as "Unable", this should correspond to the two lowest FIM® categories "Total assistance" and "Maximal assistance". The EBI category "Needs assistance by a helper with some but not all activities" was mapped by the experts to the FIM® categories "Moderate assistance" and "Minimal contact assistance". The EBI category "No direct help required, but reminders/prompts/supervision for some activities" is considered to correspond to the FIM® response category "Supervision/setup". If the highest EBI category, "Independent/independence in this activity not required (use of nappies/suprapubic catheter/indwelling catheter)", is selected, this corresponds to FIM® categories "Modified independence" and "Complete independence". An exception is made in this category however: patients who do not need to go to the toilet because they employ nappies or catheters are

mapped to the lowest ADL category, 0. It is assumed any patient who cannot walk (EBI item 6 = 0) or move from a wheelchair into bed (EBI item 5 = 0) will also require a catheter or nappies. This takes into account the different approaches of FIM® and EBI.

Table 26: Item 8: Conversion algorithm proposed by the experts

Expert workshop		
EBI 8	**ADL**	**FIM® F+J***
0	0	1 2
1	1	3 4
2	2	5
4	4 (0**)	6 7

* Variant 1A: lower of the two values for F and J
* Variant 1B: Mean score of F and J (rounded down if not a whole number)
** If EBI 5 (Bed to chair/wheelchair transfer) and EBI item 6 (Locomotion) are each scored as "0", the ADL score "4" for item 8 is revised to "0".

Item 9: Bowel management

For item pair 9, "Bowel management", the experts proposed two variants for how the FIM® and EBI response categories could be converted to a novel ADL score. As shown in Table 27, there are four response categories in EBI (category 1 does not exist), which need to be mapped to seven categories in FIM®. As is standard practice in the case of FIM®, a score in categories 1 to 5 indicates how much assistance a patient requires from another person, while a score in categories 6 and 7 indicates how well a patient can perform an activity without a helper.

Table 27: ADL item 9 – Bowel management: Comparison of EBI item 9 and FIM® item H

EBI: 9. Bowel management	
Not possible.	0
Occasional incontinence (at least once a week, but not every day); patient can change diapers independently but not clean themselves. *Or:* Bowel management requires the support of another person (at least once a week but not every day), for instance in the form of an enema.	2
Malfunctioning bowel management; however, the patient can change diapers independently, clean themselves or independently employ bowel management measures.	3
Normal bowel management (or bowel incontinence occurring less than once a week).	4

FIM: H. Bowel Management		
1	Total assist (Patient = less than 25%)	Helper
2	Maximal assist (Patient = 25% or more)	
3	Moderate assist (Patient = 50% or more)	
4	Minimal assist (Patient = 75% or more)	
5	Supervision (Patient = 100%)	
6	Modified Independence (Device)	No helper
7	Complete Independence (timely, safely)	

Variant 1: As presented in Table 28, the experts proposed for this variant that the FIM® response categories "Total assistance" and "Maximal assistance" were equivalent to EBI category "Unable". The FIM® categories "Moderate assistance" and "Minimal contact assistance" were mapped to the EBI category "Occasional episodes of incontinence". The EBI category "Impaired bowel control, patient can change nappies without assistance however, clean themselves or take measures themselves to regulate bowel movement" is considered in this variant to correspond to the FIM® categories "Supervision/setup" and "Modified independence". The EBI response "Normal bowel control" and the FIM® response "Total independence" are the corresponding highest categories.

Table 28: Item 9: Conversion algorithm proposed by the experts – *variant 1*

Expert workshop - variant 1		
EBI 9	**ADL**	**FIM® H**
0	0	1 2
2	2	3 4
3	3	5 6
4	4	7

Variant 2. As can be seen in Table 29, in variant 2 FIM® and EBI responses were mapped into broader categories, leaving only three possible scores in ADL. In this variant, the lowest EBI category ("Unable") corresponds to three FIM® categories ("Total assistance" though to "Moderate assistance"). The fourth FIM® category, "Minimal contact assistance", corresponds here to EBI category 2, "Occasional episodes of incontinence". The remaining three FIM® categories are mapped to the two highest EBI categories and receive an ADL score of 4.

Table 29: Item 9: Conversion algorithm proposed by the experts – *variant 2*

Expert workshop - variant 2		
EBI 9	**ADL**	**FIM® H**
0	0	1 2 3
2	2	4
3	4	5 6
4		7

Item 10: Bladder management

For item pair 10, "Bladder management", seven FIM® categories need to be mapped to four EBI categories. EBI category 2 does not exist (Table 30).

As can be seen in Table 31, the experts proposed the FIM® response categories "Total assistance" and "Maximal assistance" be mapped to EBI categories "Complete or very frequent incontinence (several times daily) and unable to change own nappies". The FIM® categories "Moderate assistance", "Minimal contact assistance" and "Supervision/setup" were mapped to EBI category "Incomplete incontinence". The EBI category "Complete or incomplete incontinence, requires no further assistance however" corresponds to the FIM® category "Modified independence". The best EBI response option "Normal urinary continence" and the FIM® response "Complete independence" represent the highest category in ADL.

Table 30: ADL item 10 – Bladder management: Comparison of EBI item 10 and FIM® item G

EBI: 10. Bladder management		FIM: G. Bladder Management		
Total or very frequent incontinence (numerous times per day); unable to change diapers independently. *Or:* Unable to independently operate their suprapubic cystostomy or in-dwelling catheter or cannot catheterise independently.	0	1	Total assist (Patient = less than 25%)	Helper
		2	Maximal assist (Patient = 25% or more)	
Incomplete incontinence (once a day at most); unable to change diapers or a condom catheter independently and cannot clean themselves.	1	3	Moderate assist (Patient = 50% or more)	
		4	Minimal assist (Patient = 75% or more)	
Complete or incomplete incontinence, but does not require any assistance (with changing diapers/condom catheter, personal hygiene, use of the suprapubic cystostomy or in-dwelling catheter or self-catheterisation.	3	5	Supervision (Patient = 100%)	
		6	Modified Independence (Device)	No helper
Normal bladder control.	4	7	Complete Independence (timely, safely)	

Table 31: Item 10: Conversion algorithm proposed by the experts

Expert workshop		
EBI 10	**ADL**	**FIM® G**
0	0	1 2
1	1	3 4 5
3	3	6
4	4	7

Item 11: Comprehension

For item pair 11, "Comprehension", the experts proposed two variants for how the FIM® and EBI response categories could be converted into a novel ADL score. As detailed in Table 32, the four response categories in EBI are to be mapped to seven FIM® categories. For this item, an additional item also documents whether comprehension refers to "Auditory", "Visual" or "Both".

Table 32: ADL item 11 – Comprehension: Comparison of EBI item 11 and FIM® item N

additional item	
1	Auditory
2	Visual
3	Both

EBI: 11. Cognitive comprehension	
Not possible. Cannot understand even simple instructions or questions; is unable to understand the written word or to respond reliably to prompts conveyed via facial expression or gesture.	0
Understands simple instructions (e.g. "Please take your tablets") in spoken or written form or if expressed facially or by gesture.	1
Understands complex correlations (e.g. "Please take these tablets before eating your meal"), but does not always act reliably. *Or:* Can only reliably understand written communication.	3
Normal comprehension (includes patients who depend on hearing aids, but not patients who only understand written communication).	4

FIM: N. Comprehension (akustisch /visuell)		
1	Total assist (Patient = less than 25%)	Helper
2	Maximal assist (Patient = 25% or more)	
3	Moderate assist (Patient = 50% or more)	
4	Minimal assist (Patient = 75% or more)	
5	Supervision (Patient = 100%)	
6	Modified Independence (Device)	No helper
7	Complete Independence (timely, safely)	

Variant 1. As presented in Table 33, the experts proposed for this variant that the FIM® response category "Total assistance" was equivalent to EBI category "Unable". The FIM® categories "Maximal prompting"," Moderate prompting" and "Minimal prompting" were mapped to the EBI category "Comprehends simple instructions". The EBI category "Comprehends complex facts and issues" in this variant is considered to be equivalent to the FIM® category "Standby prompting". The EBI category "Normal comprehension" and the FIM® categories "Modified independence" and "Complete independence" represent the highest category in ADL.

Table 33: Item 11: Conversion algorithm proposed by the experts – *variant 1*

Expert workshop variant 1		
EBI 11	**ADL**	**FIM® N**
0	0	1
1	1	2
		3
		4
3	3	5
4	4	6
		7

Variant 2. As can be seen in table 34, in variant 2 the FIM® response category "Total assistance" was also mapped to the EBI category "Unable". In contrast to variant 1, in this case the EBI category "Comprehends simple instructions" corresponds to the four FIM® categories "Maximal prompting" through to "Standby prompting". The sixth category in FIM®, "Modified independence", corresponds to category 3 in EBI, "Comprehends complex facts and issues". Finally, the highest FIM® category, "Complete independence", is mapped to the highest EBI category, "Normal comprehension".

Table 34: Item 11: Conversion algorithm proposed by the experts – *variant 2*

Expert workshop variant 2		
EBI 11	**ADL**	**FIM® N**
0	0	1
1	1	2
		3
		4
		5
3	3	6
4	4	7

Item 12: Expression

For item pair 12, "Expression", the experts proposed two variants for how the FIM® and EBI response categories could be converted to a novel ADL score. As detailed in Table 35, the four response categories in EBI are to be mapped to seven FIM® categories. Similar to "Comprehension", for this FIM® item an additional item documents whether the expression is "Vocal", "Nonvocal" or "Both".

Table 35: ADL item 12 – Expression: Comparison of EBI item 12 and FIM® item O

additional item	
1	Vocal
2	Nonvocal
3	Both

EBI: 12. Expression	
Expression is never or very rarely possible.	0
Can only express simple daily concerns (e.g. hunger, thirst, etc.), be it with or without an aid (e.g. in writing or via a communicator).	1
Is able to express virtually everything, but only with aids (e.g. in writing, with communicator).	3
Is able to express practically everything without using aids (although permitted are grammatical errors, slight difficulty finding the right words or slightly unclear speech).	4

FIM: O. Expression		
1	Total assist (Patient = less than 25%)	Helper
2	Maximal assist (Patient = 25% or more)	
3	Moderate assist (Patient = 50% or more)	
4	Minimal assist (Patient = 75% or more)	
5	Supervision (Patient = 100%)	
6	Modified Independence (Device)	No helper
7	Complete Independence (timely, safely)	

Variant 1. As shown in Table 36, the experts proposed for this variant that the FIM® response category "Total assistance" was equivalent to EBI category "Can never, or almost never, express him-/herself". The four FIM® categories "Maximal", "Moderate", "Minimal prompting" and "Standby prompting" were mapped to EBI category "Can only express simple, every-day facts and issues (e.g. hunger, thirst)". The EBI category "Can express him-/herself regarding almost everything, but only using auxiliary aids" is considered equivalent to the FIM® category "Modified independence". Finally, the highest EBI category "Can express him-/herself regarding almost everything without using aids" and the highest FIM® category "Complete independence" were mapped together.

Table 36: Item 12: Conversion algorithm proposed by the experts – *variant 1*

Expert workshop variant 1		
EBI 12	**ADL**	**FIM® O**
0	0	1
1	1	2 3 4 5
3	3	6
4	4	7

Variant 2. As can be seen in Table 37, in variant 2 FIM® and EBI responses were mapped in broader categories, leaving only three possible scores in ADL. In this variant, the lowest EBI category, "Can almost never express him-/herself", corresponds to the two FIM® categories "Total assistance" and "Maximal prompting". The FIM® categories "Moderate prompting", "Minimal prompting" and "Standby prompting" correspond to EBI category 1, "Can only express simple, every-day facts and issues (e.g. hunger, thirst)". The remaining two FIM® categories are mapped to the two highest EBI categories and receive an ADL score of 4.

Table 37: Item 12: Conversion algorithm proposed by the experts – *variant 2*

Expert workshop variant 2		
EBI 12	**ADL**	**FIM® O**
0	0	1 2
1	1	3 4 5
3 4	4	6 7

Item 13: Social interaction

There are only three EBI levels of function for item pair 13 "Social interaction"; "Is always or almost always uncooperative", "Is occasionally uncooperative, distanced or withdrawn" and "Normal social interaction" (Table 38).

Table 38: ADL item 13 – Social interaction: Comparison of EBI item 13 and FIM® item P

EBI: 13. Social interaction	
Always or largely uncooperative (e.g. resists efforts made by nursing staff), aggressive, unresponsive or reclusive.	0
Occasionally uncooperative, aggressive, unresponsive or reclusive	2
Normal social interaction.	4

FIM: P. Social Interaction		
1	Total assist (Patient = less than 25%)	Helper
2	Maximal assist (Patient = 25% or more)	
3	Moderate assist (Patient = 50% or more)	
4	Minimal assist (Patient = 75% or more)	
5	Supervision (Patient = 100%)	
6	Modified Independence (Device)	No helper
7	Complete Independence (timely, safely)	

As presented in Table 39, the experts proposed that the FIM® response categories "Total assistance" and "Maximal direction" were equivalent to EBI category "Is always or almost always uncooperative". The FIM® categories "Moderate direction", "Minimal direction" and "Supervision" were mapped to EBI category "Is occasionally uncooperative, distanced or withdrawn". The highest EBI category "Normal social interaction" corresponds to the two FIM® categories "Modified independence" and "Complete independence".

Table 39: Item 13: Conversion algorithm proposed by the experts

Expert workshop		
EBI 13	**ADL**	**FIM® P**
0	0	1 2
2	2	3 4 5
4	4	6 7

Item 14: Problem solving

Similar to the previous item pair, there are only three EBI levels of function for item pair 14, "Needs considerable assistance due to above-named disorders", "Needs minimal assistance due to above-named disorders" and "Needs no assistance despite above-named disorders". Examples of everyday disorders in problem solving are detailed in Table 40.

Table 40: ADL item 14 – Problem solving: Comparison of EBI item 14 and FIM® item Q

EBI: 14. Problem solving	
Examples of difficulty with solving everyday problems include: overhasty actions, (e.g. getting out of the wheelchair without first applying the brakes), inflexible behaviour (e.g. difficulty adjusting to a change in the daily routine), failure to keep appointments, difficulty with self-medication (that is not caused by motor disability), inadequate understanding of disabilities and their consequences in daily life .	
Requires significant assistance on account of any of the problems listed above.	0
Requires a little assistance on account of any of the problems listed above.	2
Requires no assistance on account of any of the problems listed above.	4

FIM: Q. Problem Solving		
1	Total assist (Patient = less than 25%)	Helper
2	Maximal assist (Patient = 25% or more)	
3	Moderate assist (Patient = 50% or more)	
4	Minimal assist (Patient = 75% or more)	
5	Supervision (Patient = 100%)	
6	Modified Independence (Device)	No helper
7	Complete Independence (timely, safely)	

As presented in Table 41, the experts proposed that the FIM® response categories "Total assistance" and "Maximal direction" were equivalent to EBI category "Needs considerable assistance due to above-named disorders". The FIM® categories "Moderate direction", "Minimal direction" and "Supervision" were mapped to EBI category "Needs minimal assistance due to above-named disorders". The highest EBI category "Needs no assistance despite above-named disorders" corresponds to the two FIM® categories "Modified independence" and "Complete independence". The mapping for this item pair was the same as for the previous item pair.

Table 41: Item 14: Conversion algorithm proposed by the experts

Expert workshop		
EBI 14	**ADL**	**FIM® Q**
0	0	1 2
2	2	3 4 5
4	4	6 7

Item 15: Memory/learning capability/orientation

For item pair 15, "Memory/learning capability/orientation", seven FIM® categories need to be mapped to five EBI categories (Table 42).

Table 42: ADL item 15 – Memory/learning capability/orientation: Comparison of EBI item 15 and FIM® item R

EBI: 15. Memory/learning capacity/orientation	
Patient is disoriented or confused with a strong disposition to flight.	0
Patient is disoriented or confused but has no disposition to flight; however, patient does have difficulty finding their way around the hospital *Or:* Completely incapable of retaining new information (e.g. cannot remember their contact person at the hospital even after a number of meetings, forgets what has been discussed or agreed or where items are stored) and cannot use external memory aids (e.g. calendars, notepads).	1
Often needs prompting.	2
Sometimes needs prompting.	3
No impairment of relevance to everyday life. *Or*: Can use external memory aids effectively. *Or*: Despite memory or orientation impairment, the patient does not need any additional (nursing) care (e.g. completely immobile patients with severe orientation difficulties).	4

FIM: R. Memory		
1	Total assist (Patient = less than 25%)	Helper
2	Maximal assist (Patient = 25% or more)	
3	Moderate assist (Patient = 50% or more)	
4	Minimal assist (Patient = 75% or more)	
5	Supervision (Patient = 100%)	
6	Modified Independence (Device)	No helper
7	Complete Independence (timely, safely)	

As shown in Table 43, the experts proposed the FIM® response categories "Total assistance" and "Maximal prompting" correspond to the EBI categories "Is disorientated or confused and is highly likely to stray" and "Is completely unable to retain new information", with an ADL score of 0. The FIM® categories "Moderate prompting", "Minimal prompting" and "Standby prompting" were mapped to the two EBI

categories "Needs frequent reminders" and "Only needs occasional reminders". The highest EBI category "No impairment relevant to daily routine" corresponds to the two highest FIM® categories "Modified independence" and "Complete independence". The number of categories in EBI are reduced in ADL from 5 to 3 for this item, to appropriately map the FIM® categories.

Table 43: Item 15: Conversion algorithm proposed by the experts

Expert workshop		
EBI 15	**ADL**	**FIM® R**
0	0	1
1		2
2	2	3
3		4
		5
4	4	6
		7

3.2. Validation of ADL scores

3.2.1. Description of the sample population

The required number of 31 to 70 cases per clinic with completed FIM® and EBI records was achieved (Figure 2). To interpret the case number figures, it should be noted that clinics 1 and 3 represent two sites within one group of clinics. The mean age was 67.2 years, with a range of 18 to 92 years (Figure 3). The mean age of patients of the individual five clinics involved was 64 to 73 years (Figure 61). 50.8% of the study patients were male (Figure 4). The percentage of males in the individual clinics ranged from 43% to 60% (Figure 62). 90.0% of patients were Swiss nationals (Figure 5); this figure for the individual clinics ranged from 81% to 100% of patients (Figure 63). The score for mean comorbidity as measured using CIRS was 13.2 points (Figure 6). Comorbidity scores in the individual clinics involved ranged from 10 to 20 points (Figure 64). Individual percentages for the five standardised impairment categories in the whole sample population ranged from 15.8% to 26.0% for EBI (Figure 7) and from 16.2% to 24.2% for FIM® (Figure 8). With the exception of one clinic, this distribution was approximately mirrored in the individual clinics (Figures 65 and 66). The sample population parameters for the individual clinics are presented in graph form in Appendix A1.

Figure 2: No. of cases per clinic

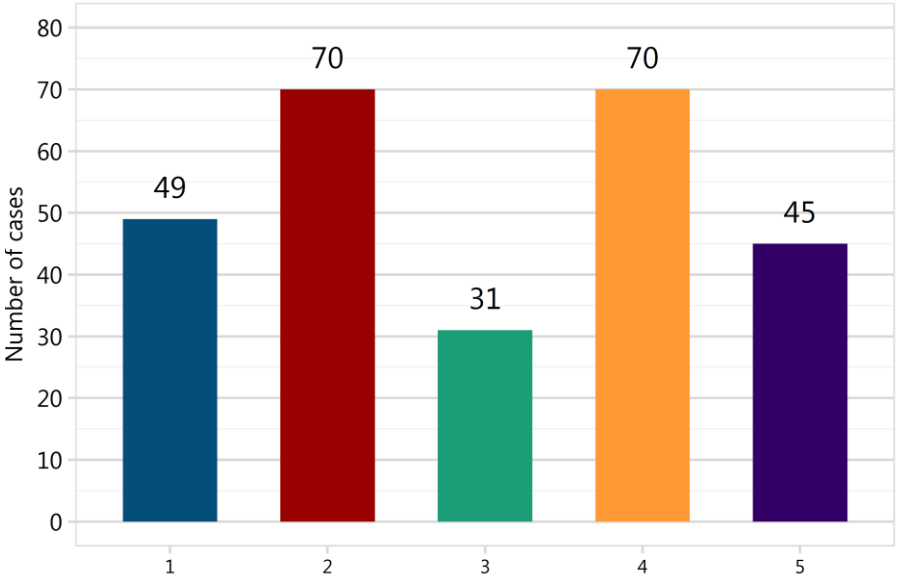


Figure 3: Age

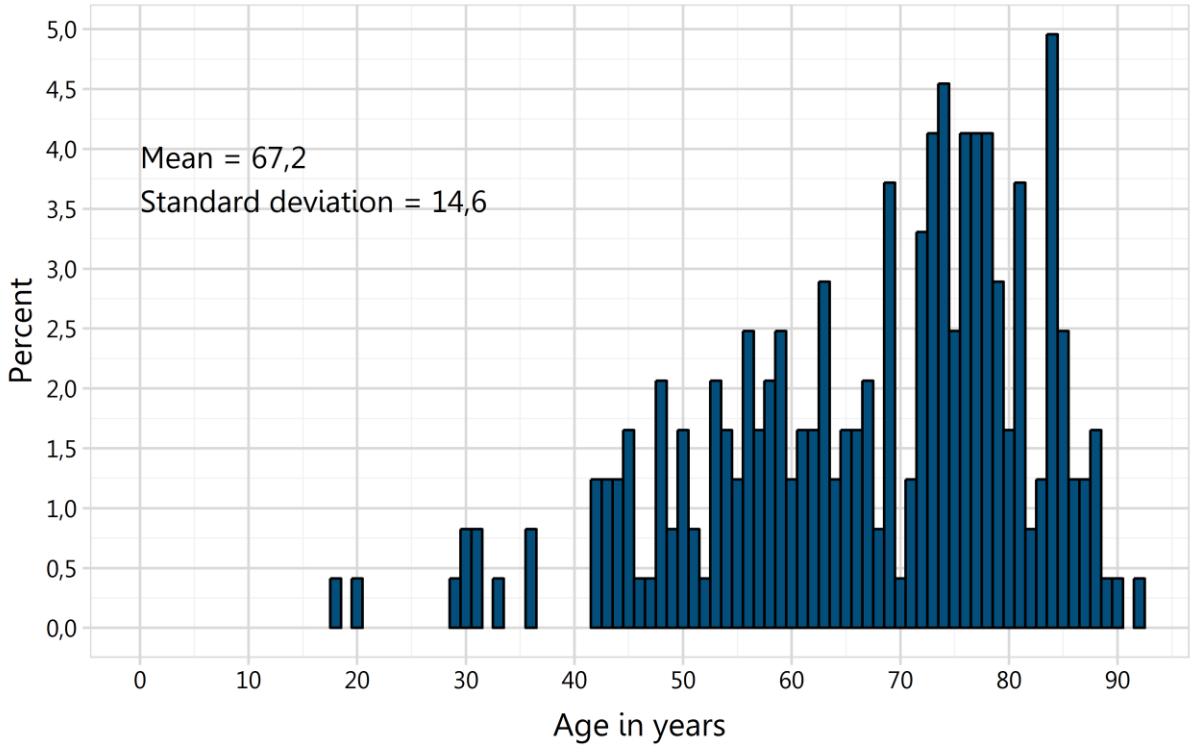


Figure 4: Gender

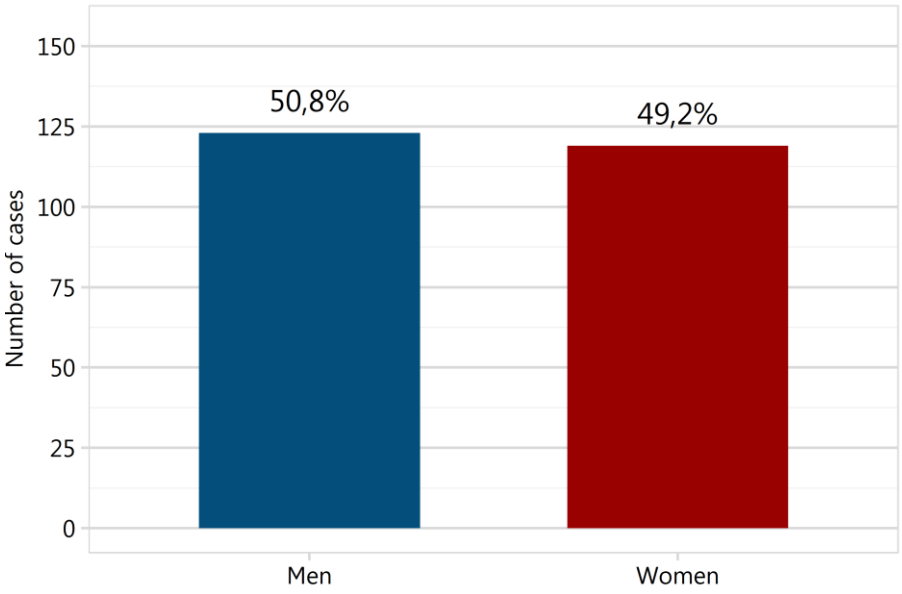


Figure 5: Nationality

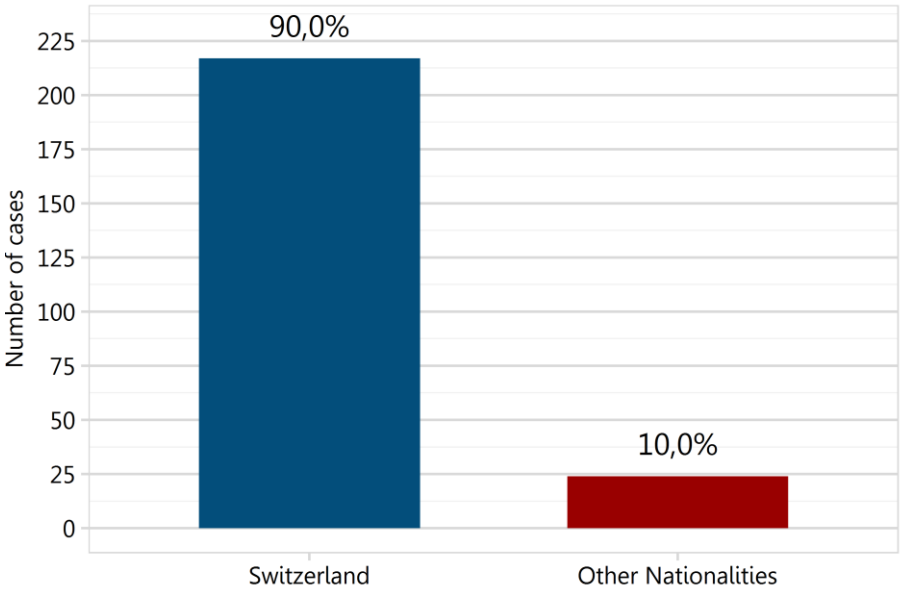


Figure 6: Comorbidity (CIRS)

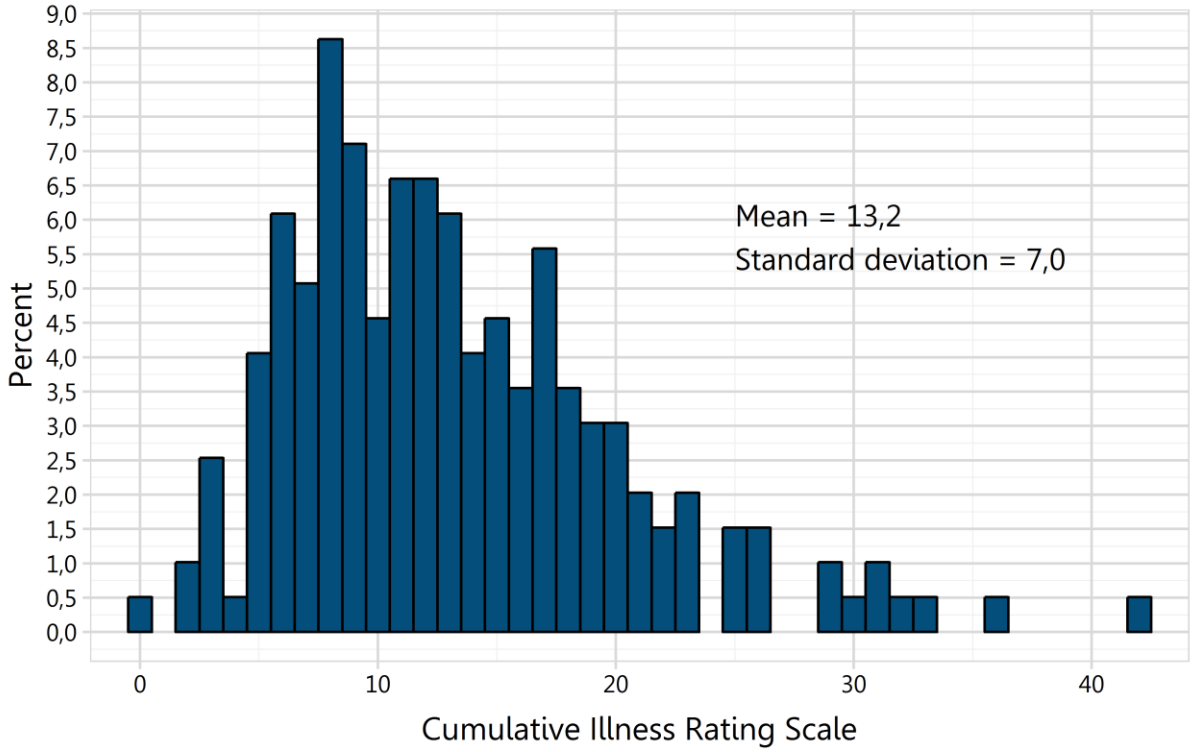


Figure 7: EBI categories

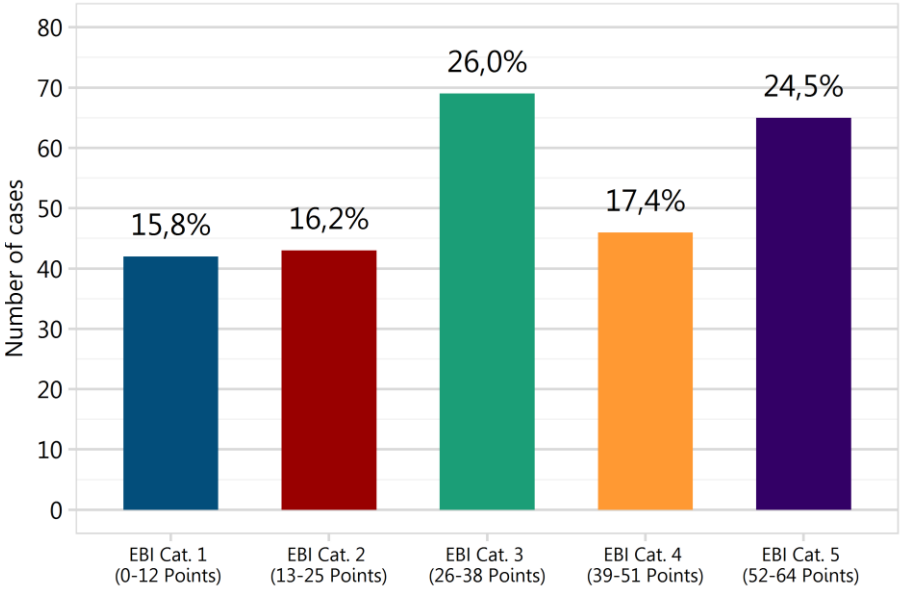
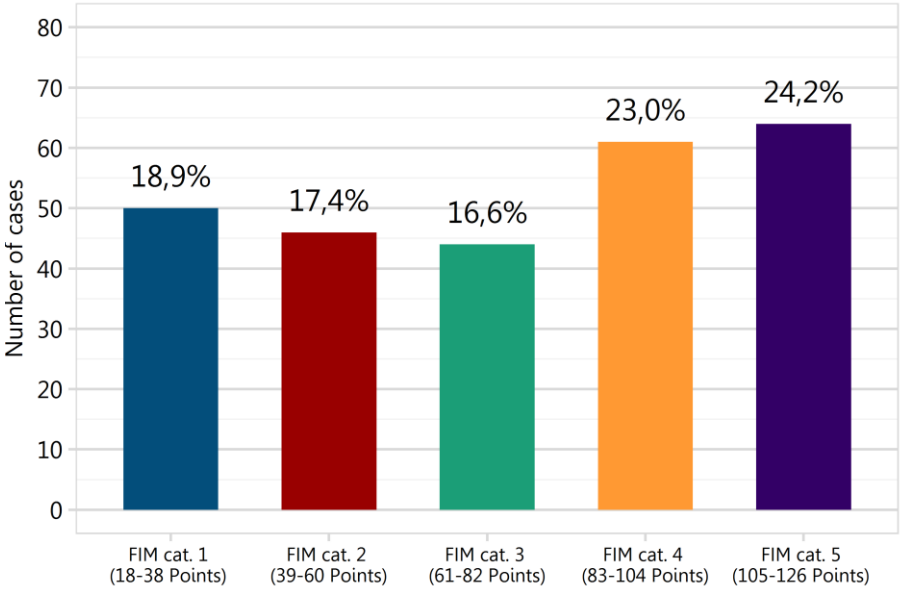


Figure 8: FIM® categories



3.2.2. Analysis at the item level

Item 1: Eating and drinking

The transition plot for item pair 1 shows that the response categories 4 and 2 in EBI, and 5 in FIM® were the most frequently selected (Figure 9). The FIM® categories 2 and 4 and the EBI category 3 were selected by scorers somewhat less frequently, each with a maximum of 6% of patients.

The arrows indicate that it is fairly clear that FIM® categories 1 and 3–7 can be mapped to corresponding EBI categories. For most categories, the content-related mappings of FIM® and EBI categories closely agree with the empirical data. In contrast with both expert workshop proposals, FIM® category 6 generally corresponds to EBI category 4. FIM® category 2 is selected approximately as often with EBI category 0 as it is with EBI 2. In the expert workshop the mapping proposal for both variants was to EBI category 0.

Figure 9: Transition plot for item 1 (Eating and drinking)

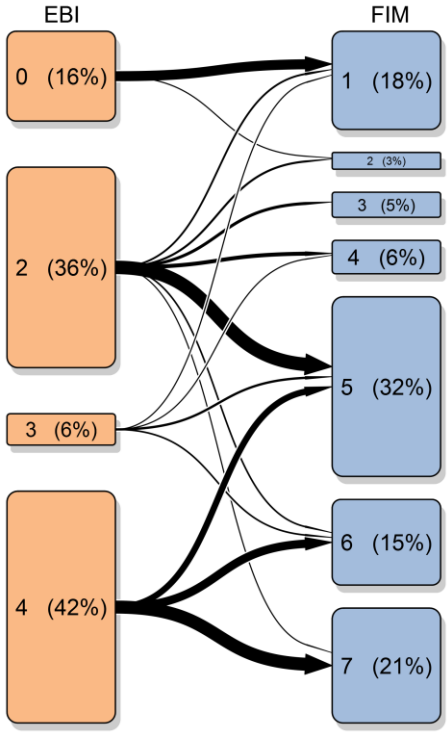


Figure 10 presents the implementation of the ADL algorithm in both expert variants for this item. The ADL score as calculated on the basis of FIM® (y-axis) is plotted against the ADL score on the basis of EBI (x-axis); these should ideally be identical. Spearman's rank correlation coefficient shows there is close correlation, and the kappa values indicate close levels of agreement. In variant 1, the difficulty in agreeing on the EBI correspondence with ADL in category 4, in particular, can be seen, which corresponds in the case of FIM® to ADL scores of 2, 3 or 4. Implementation of the second variant shows that the ADL

score of 0 calculated on the basis of FIM® partially corresponds to an EBI score of 2. This can be attributed to the – empirically not justifiable – mapping of FIM® categories 3 and 4 to EBI category 0.

As a result, it was proposed that both variants be mixed. The result of mapping on the basis of the empirical data largely conforms to variant 1. The exception is that, as in variant 2, FIM® category 6 is mapped to EBI category 4 (instead of 3). This significantly increases the level of agreement per the kappa value.

Figure 10: Scatter plot for item 1 (Eating and drinking) based on expert workshop *variants 1 and 2*

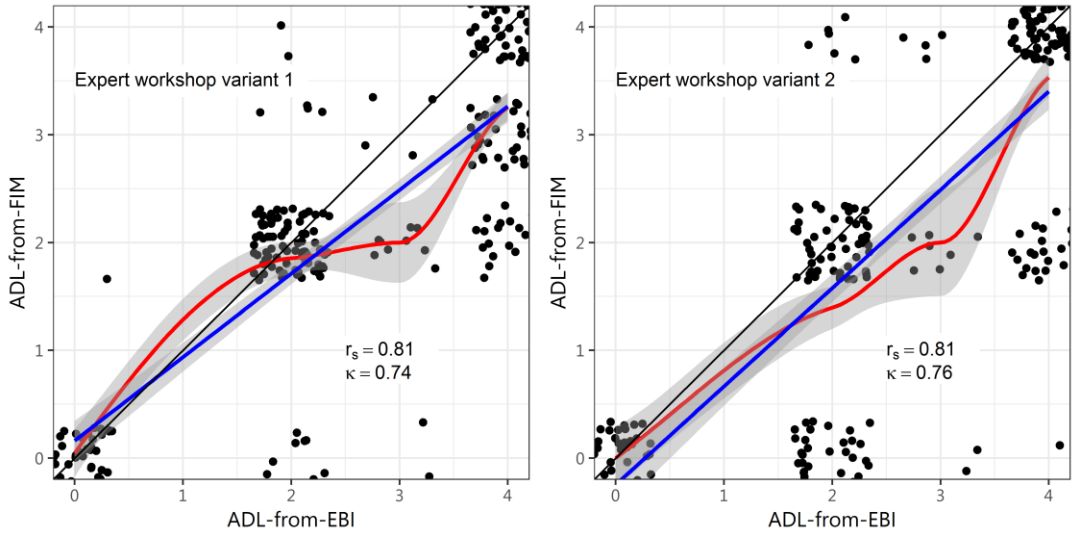
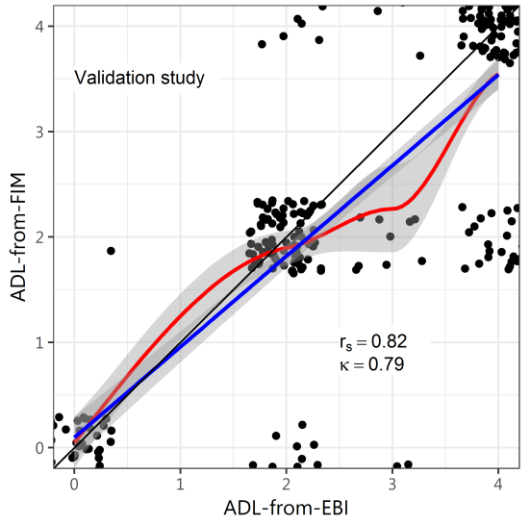


Figure 11: Scatter plot for validation study for item 1 (Eating and drinking)

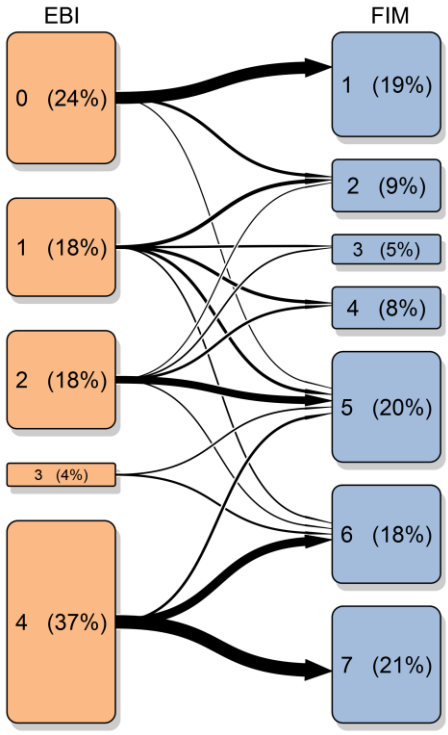


Validation study		
EBI 1	**ADL**	**FIM® A**
0	0	1 2
2	2	3 4 5
3	3	
4	4	6,7

Item 2: Personal hygiene

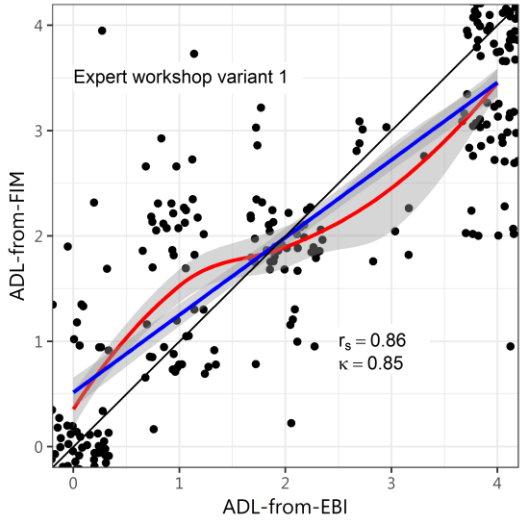
The transition plot for item pair 2 shows that the response categories 4 (37%) and 0 (24%) in EBI, and 7 (21%), 5 (20%), 1 (19%) and 6 (18%) in FIM® were the most frequently selected (Figure 12). The thickness of the arrows indicates which combinations of EBI and FIM® scores occurred most frequently. it is clear that EBI category 0 was selected most frequently together with FIM® category 1. This was also the case for the score pairs EBI 2 and FIM® 5, EBI 4 and FIM® 6, and EBI 4 and FIM® 7. All other score combinations are less clear. In particular, FIM® categories 2, 3 and 4 occur more or less equally in combination with EBI categories 0, 1 and 2. This uncertainty in mapping is also mirrored by the three variants proposed by the experts, in dealing with precisely these three score pairs.

Figure 12: Transition plot for item 2 (Personal hygiene)

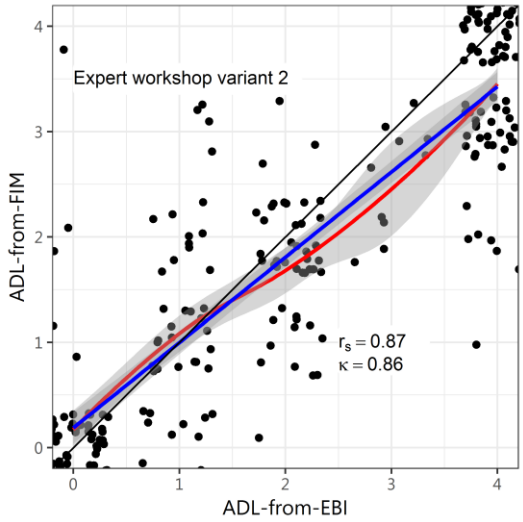


In Figure 13 the empirical mapping of the three expert variants to produce an ADL score is shown in three scatter plots. In particular, in ADL-from-EBI category 2 the scores are scattered in ADL-from-FIM® in all three variants. The key figures for all three variants are very similar and indicate a very high level of correspondence between the ADL scores from FIM® and from EBI. The third variant, in which EBI categories 3 and 4 are jointly mapped to an ADL score of 4, is less than optimally related at the lower end (ADL 0 to 2); the red regression curve obtained using the LOWESS method is relatively far from the ideal relationship represented by the straight black line.

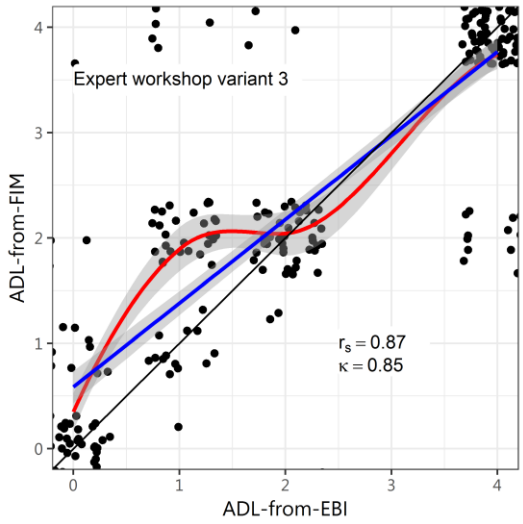
Figure 13: Scatter plots for item 2 from expert workshop variants 1–3



Expert workshop variant 1		
EBI 2	**ADL**	**FIM® B**
0	0	1
1	1	2
		3
2	2	4
		5
3	3	6
4	4	7



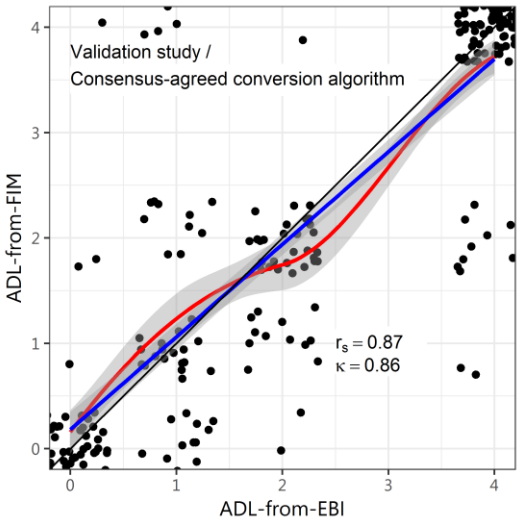
Expert workshop variant 2		
EBI 2	**ADL**	**FIM® B**
0	0	1
		2
1	1	3
		4
2	2	5
3	3	6
4	4	7



Expert workshop variant 3		
EBI 2	**ADL**	**FIM® B**
0	0	1
1	1	2
2	2	3
		4
		5
3	4	6
4		7

As a result, it was proposed that variants 2 and 3 be mixed. The mapping for ADL 0, 1 and 2 corresponds to variant 2, there is no ADL 3 score (as in variant 3), and an EBI score of 3 is mapped to give an ADL score of 4 (Figure 14). The results for Spearman's rank correlation coefficient and kappa are roughly equivalent to those for the second expert variant.

Figure 14: Scatter plots for item 2 from the validation study



Validation study		
EBI 2	**ADL**	**FIM® B**
0	0	1
		2
1	1	3
		4
2	2	5
3	4	6
4		7

Item 3: Dressing and undressing

For item pair 3, "Dressing and undressing", two transition plots were generated (Figure 15). The first transition plot (labelled a, on the left-hand side) shows the percentage distribution of selected EBI and FIM® scores when the lower of the two FIM® scores from items D (Dressing upper body) and E (Dressing lower body) were chosen. In this case, a patient who can independently put on a pair of trousers but not a t-shirt might receive a very low total FIM® score. The second transition plot (labelled b, on the right-hand side) shows the percentage distribution of selected EBI and FIM® scores when the mean of the two FIM® scores from items D and E was recorded and then rounded down. In this case, a patient with the same level of severity of impairment (trousers can be put on, t-shirt can't) would receive, in total, a higher overall FIM® score. In the percentage distributions it is apparent that the lowest FIM® category occurs more often (35% of cases) in the first variant (a, lower score is chosen) than in the variant in which the mean score is taken (29% of cases). In the mean score version (b), FIM® scores 3, 4 and 5 are cited more frequently. In the variant in which the lower of the two scores is chosen, these scores are more frequently assigned in the form of the lower or the lowest categories.

Figure 15: Transition plot for item 3 (Dressing and undressing): a) Lower score of the FIM® items D and E; b) Mean score rounded down for FIM® items D and E

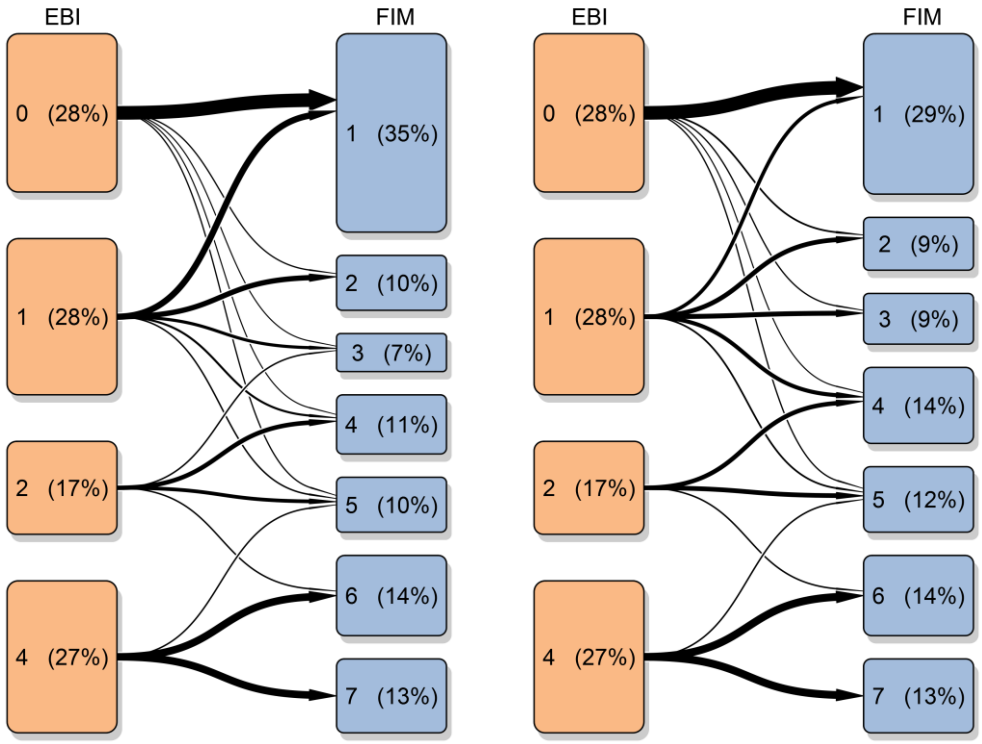
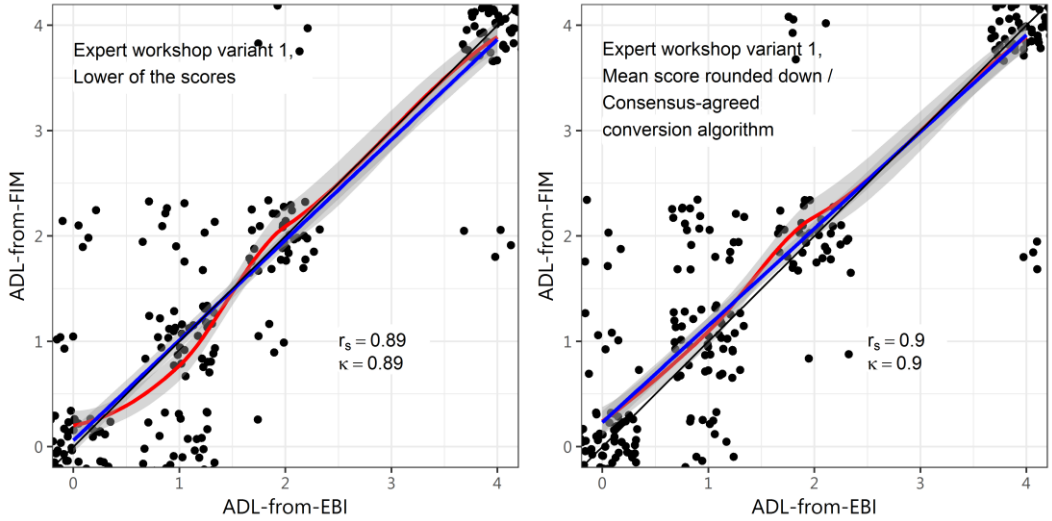


Figure 16 shows the empirical mapping of the two versions (a – lower score, b – mean score rounded down) for the mapping variant 1 by the experts. Figure 17 shows the empirical mapping of the two versions (a – lower score, b – mean score rounded down) for the mapping variant 2 by the experts. Not only the key values but also the regression curves generated clearly indicate that the first mapping variant by the experts is the variant that fits better empirically. The correlation according to Spearman's rank correlation coefficient is closer in the case of the second variant (in which a score of 0 in EBI and in ADL has no equivalent in FIM®, but in this variant the kappa values (0.78 and 0.74) are much lower than in the first variant (0.89 and 0.90).

In the first variant, the version in which the mean score rounded down of FIM® items D and E was employed has marginally better agreement parameters than the version in which the lower of the two scores was taken.

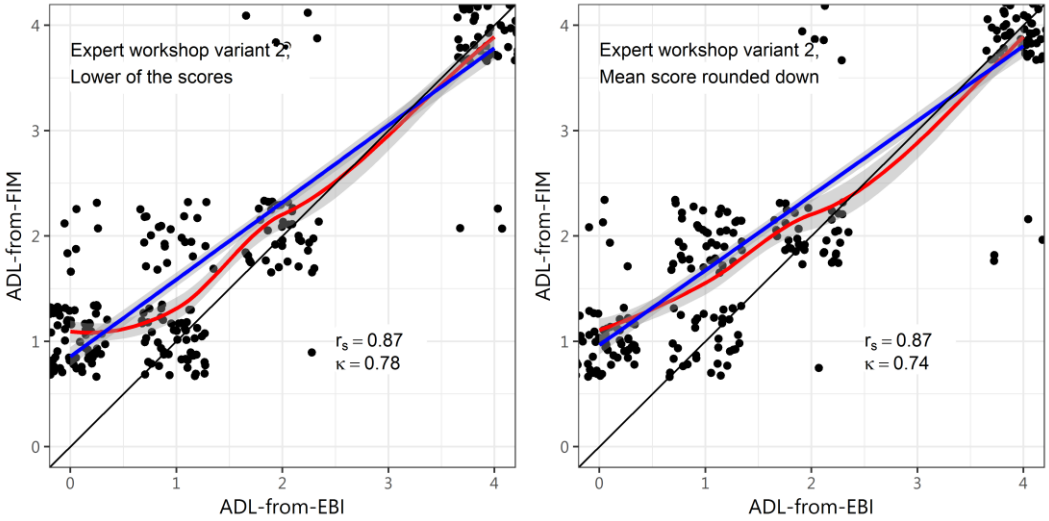
Figure 16: Scatter plots for item 3 from expert workshop *variant 1*



Expert workshop variant 1		
EBI 3	**ADL**	**FIM® D+E***
0	0	1
1	1	2
		3
2	2	4
		5
4	4	6
		7

* Variant 1A: lower of the scores for D and E
* Variant 1B: mean score of D and E (rounded down, if not a whole number)

Figure 17: Scatter plots for item 3 from expert workshop *variant 2*



Expert workshop variant 2		
EBI 3	**ADL**	**FIM® D+E***
0	0	
1	1	1,2
2	2	3 4 5
4	4	6 7

* Variant 2A: lower of the scores for D and E
* Variant 2B: mean score of D and E (rounded down if not a whole number)

Item 4: Bathing/showering/taking a complete sponge bath

For item pair 4, "Bathing/showering/taking a complete sponge bath", two transition plots were generated (Figure 18). The first transition plot (labelled a, on the left-hand side) shows the percentage distribution of selected EBI and FIM® scores when the lower of the two FIM® scores from items C (Bathing/showering/taking a complete sponge bath) and K (Transfer to shower/tub) were chosen. In this case, a patient who can independently take a shower but not get into the shower might receive a very low total FIM® score. The second transition plot (labelled b, on the right-hand side) shows the percentage distribution of selected EBI and FIM® scores when the mean of the two FIM® scores from items C and K was taken and then rounded down. In this case, a patient with the same level of severity of impairment (patient can shower alone but not get into the shower) would receive, in total, a higher overall FIM® score. In the case of the percentage distributions it is apparent that the lowest FIM® category occurs

more often (34% of cases) in the first variant (a, lower score is chosen) than in the variant in which the mean score is taken (31% of cases). There are only small differences in the other FIM® categories.

Figure 18: Transition plot for item 3 (Bathing/showering/taking a complete sponge bath): a) Lower score of the FIM® items C and K; b) Mean score rounded down for FIM® items C and K

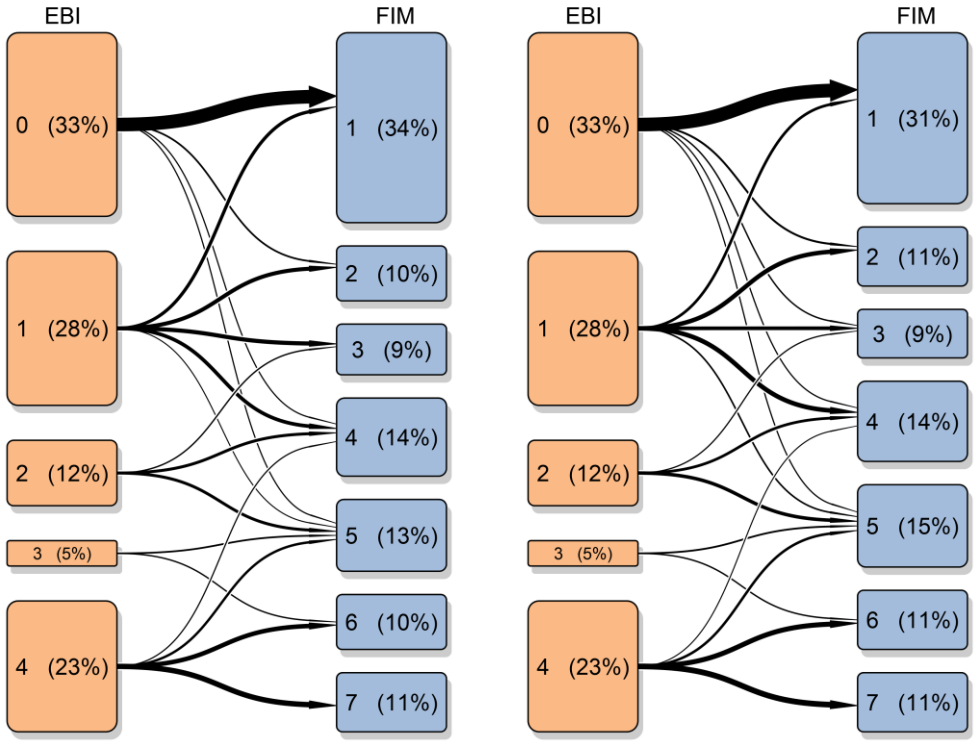
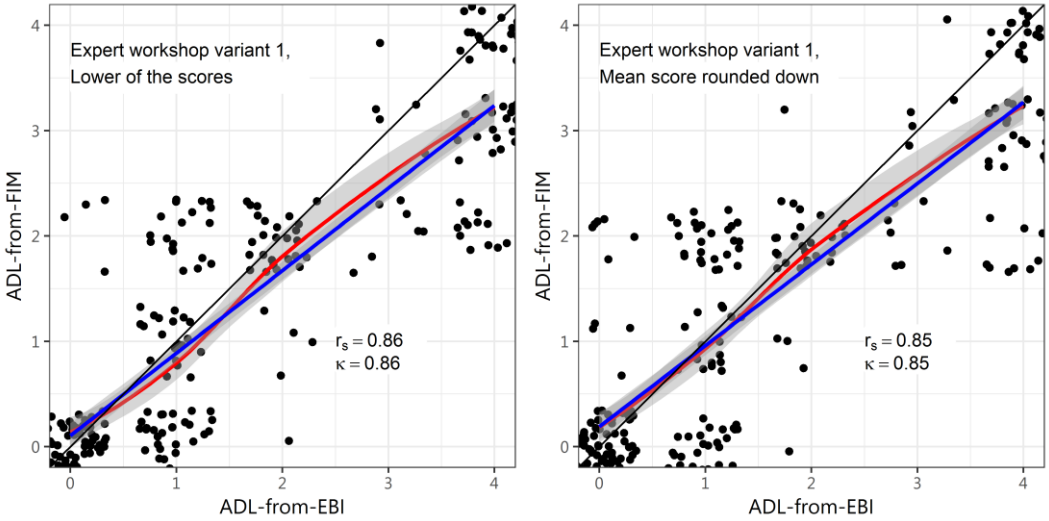


Figure 19 shows the empirical mapping of the two versions (a – lower score, b – mean score rounded down) for the *first* mapping variant by the experts. Figure 20 shows the empirical mapping of the two versions (a – lower score, b – mean score rounded down) for the *second* mapping variant by the experts. Figure 21 shows the empirical mapping of the two versions in the mapping variants from the validation study. All versions have fairly high levels of agreement, showing no major discrepancies between the ADL values-from-EBI and the ADL values-from-FIM®. Empirically the best agreement is the mapping proposal arising from the validation study, in which the lower of the two scores for FIM® items C and K was employed (version a).

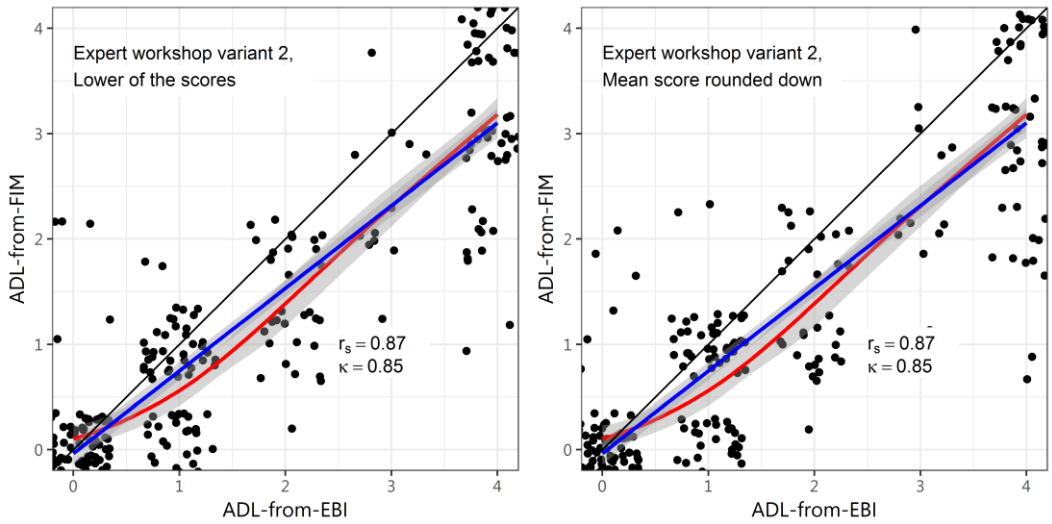
Figure 19: Scatter plots for item 4 from expert workshop *variant 1*



Expert workshop variant 1		
EBI 4	**ADL**	**FIM® C+K***
0	0	1 2
1	1	3
2	2	4 5
3	3	6
4	4	7

* Variant 1A: lower of the scores for C and K
* Variant 1B: mean score of C and K (rounded down if not a whole number)

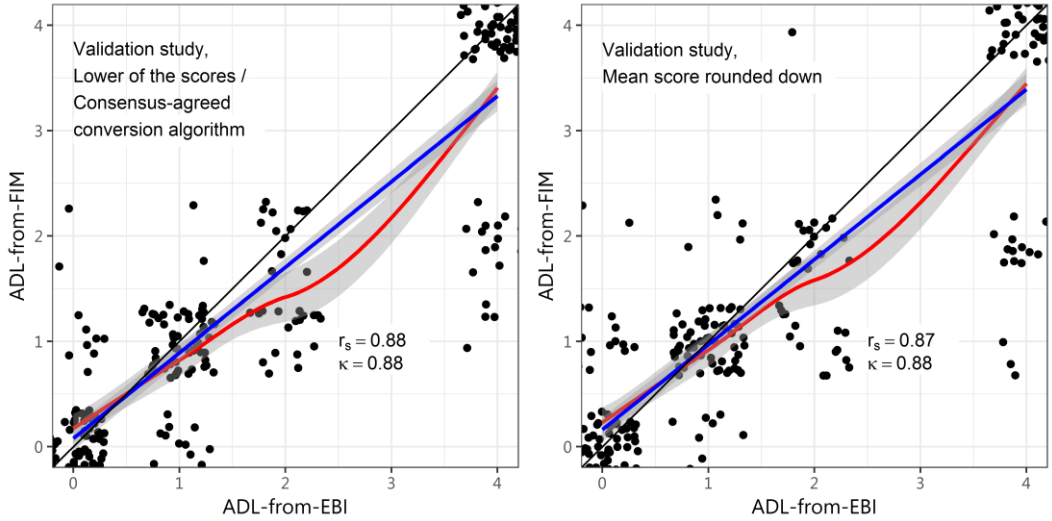
Figure 20: Scatter plots for item 4 from expert workshop *variant 2*



Expert workshop variant 2		
EBI 4	**ADL**	**FIM® C+K***
0	0	1 2
1	1	3 4
2	2	5
3	3	6
4	4	7

* Variant 2A: lower of the scores for C and K
* Variant 2B: mean score of C and K (rounded down if not a whole number)

Figure 21: Scatter plots for item 4 from the validation study



Validation study		
EBI 4	**ADL**	**FIM® C+K***
0	0	1
1	1	2
		3
		4
2	2	5
3	4	6
4		7

*Variant VA: lower of the scores for D and E
*Variant VB: mean score of D and E (rounded down if not a whole number)

Item 5: Bed to chair/wheelchair transfer

The transition plot for item pair 5 shows that response categories 0 and 4 in EBI, and 1, 6 and 7 in FIM® were the most frequently selected (Figure 9). In most cases, EBI category 0 corresponds with FIM® category 1 and EBI category 4 with FIM® categories 6 and 7. The mid-level categories cannot be mapped so readily. Overall, the level of agreement between the content-related mapping by the experts and the empirical data is high.

Figure 22: Transition plot for item 5 (Bed to chair/wheelchair transfer)

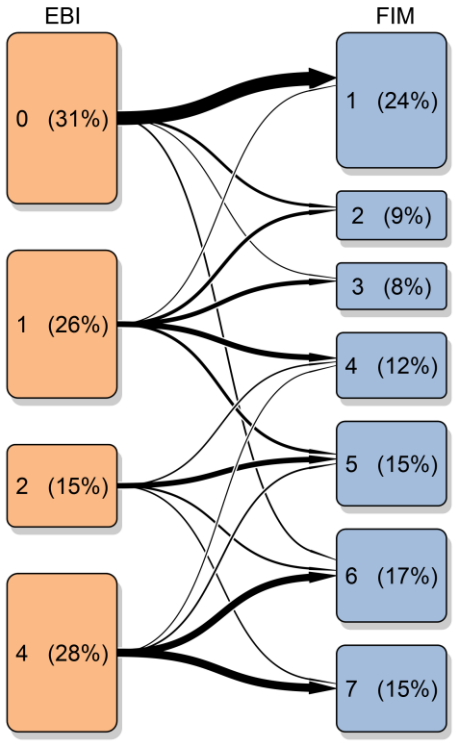
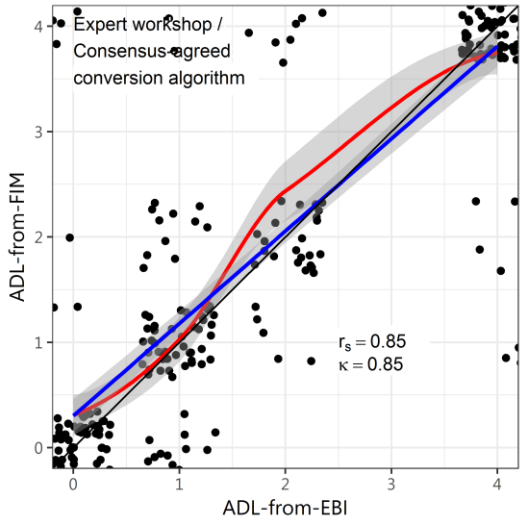


Figure 23 shows the implementation of the ADL algorithm by the experts for this item. Spearman's rank correlation coefficient shows there is close correlation, and the kappa values indicate close levels of agreement. The proposal to map categories 1 and 2 in FIM® to EBI category 0 did not fit better empirically than the original proposal (Figure 71).

Figure 23: Scatter plots for item 5 from expert workshop



Expert workshop		
EBI 5	**ADL**	**FIM® I**
0	0	1 2
1	1	3 4
2	2	5
4	4	6 7

Item 6: Locomotion

The transition plot for item pair 6 shows that response categories 0 and 1 in EBI, and 1 and 5 in FIM® were the most frequently selected (Figure 24). In most cases EBI category 0 corresponds with FIM® category 1. In all categories the empirical mapping is not particularly clear. EBI category 1, in particular, is a case in point, as it was selected at the same time as all of the FIM® categories. However, this ambiguity is not as large as a first glance at the transition plot might suggest, as the mapping is also subdivided into the response categories of the additional items. The expert-proposed mapping is described in detail above in section 3.1 ("item 6"). In summary, if patients use a wheelchair to move, the EBI categories 0 to 2 are sufficient to document an ADL score, while in FIM® this is done by selecting categories 1 to 6.

Figure 25 shows as a scatter plot the empirical mapping of the expert variant to produce an ADL score. The scores are widely scattered, in particular for EBI category 1. The regression curves are an indication that the ADL score agreement is somewhat less than perfect (ideal indicated by the black diagonal line).

In view of this empirical data, a still further differentiated mapping of FIM® and EBI with regards to the proposed additional item was undertaken. As a result, there are individual conversion algorithms for patients who can walk, for patients who use a wheelchair to move and for patients who both walk and use a wheelchair. Figure 26 show the empirical ADL score distribution calculated based on the proposal from the validation study. In Figure 26 the three tables below the scatter plot detail the process. In contrast to the first expert proposal, the highest score for a patient only using a wheelchair alone is 1 (expert proposal 2). For patients assigned to the additional category "Both", who both walk and use a wheelchair, the highest score is 2, and for patients who can walk, the highest score is 4.

Figure 24: Transition plot for item 6 (Locomotion)

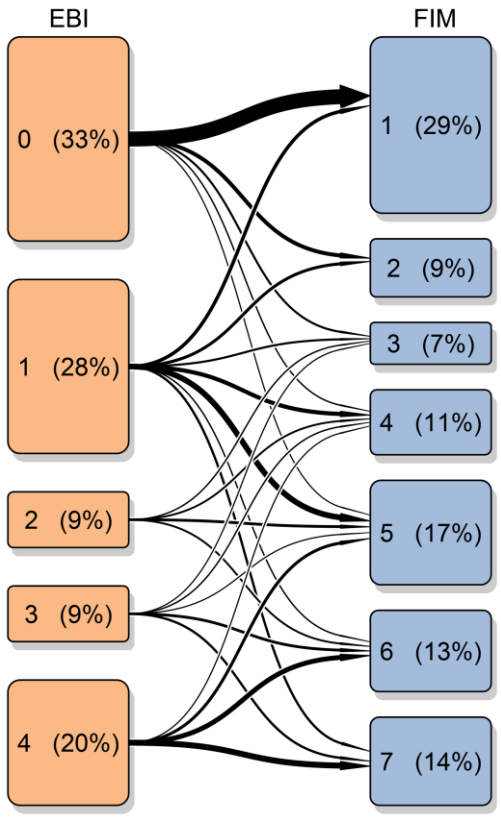


Figure 25: Scatter plot for item 6 from expert workshop

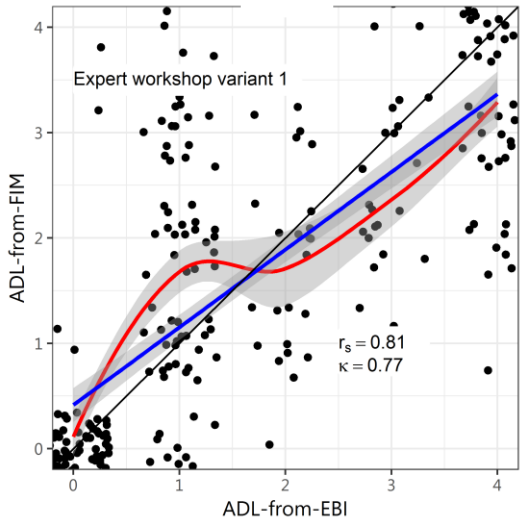
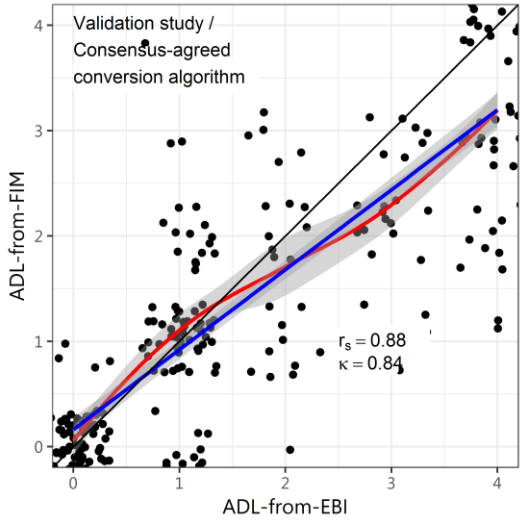


Figure 26: Scatter plot for item 6 from the validation study



Expert workshop variant Walking		
EBI 6	Gehen	FIM® L
0	0	1
		2
1	1	3
		4
2	2	5
3	3	6
4	4	7

Validation Study Wheelchair		
EBI 6	ADL	FIM® L
0	0	1
		2
1		3
		4
2	1	5
3		6
4		7

Validation Study Walking AND Wheelchair		
EBI 6	ADL	FIM® L
0	0	1
		2
1	1	3
		4
2	2	5
3		6
4		7

Item 7: Stairs

The transition plot for item pair 7 "Stairs" shows that, in contrast to most other items, most often selected are EBI category 0 (67% of cases) and FIM® category 1 (68% of cases) (Figure 27).

Figure 28 shows the empirical mapping of the expert variant to an ADL score in the form of a scatter plot. The scores are sometimes widely scattered, in particular for ADL-from-EBI category 4. In general, as the transition plot shows, most of the scores are consistent for FIM® and EBI in the lowest ADL category 0; in other words, most patients are not able to use stairs. This explains the high level of agreement.

Figure 27: Transition plot for item 7 (Stairs)

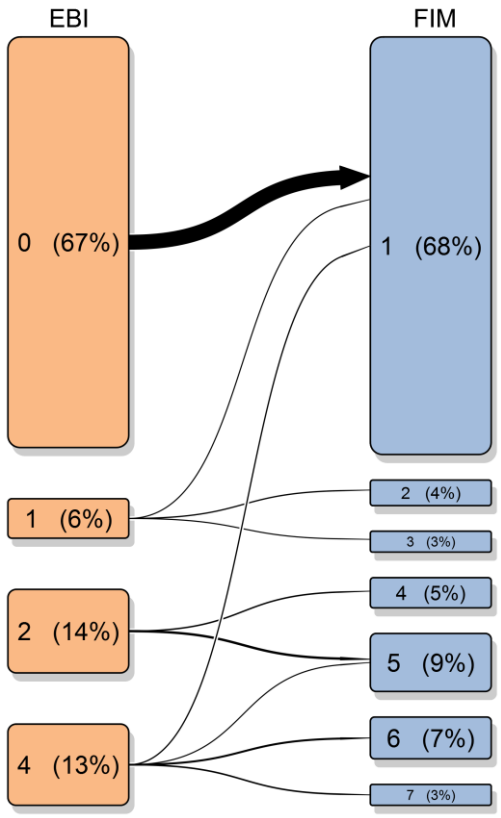
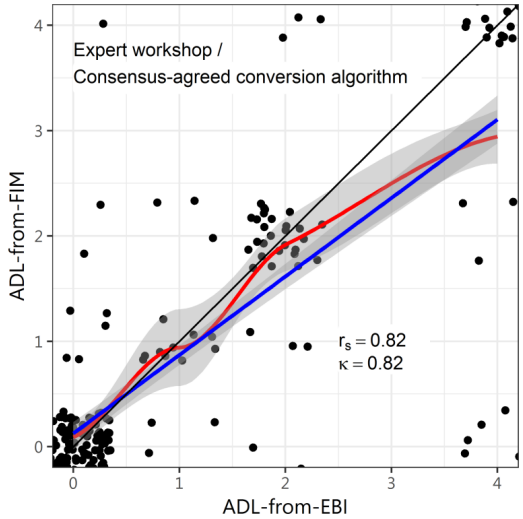


Figure 28: Scatter plot for item 7 from expert workshop



Expert workshop		
EBI 7	**ADL**	**FIM® M**
0	0	1
1	1	2 3
2	2	4 5
4	4	6 7

Item 8: Toilet use

The transition plot for item pair 8 "Toilet use" shows that in EBI most often selected are categories 4 (40% of cases), followed by 0 (28%) and 1 (25%) (Figure 29). In FIM® the empiric mappings slightly differ – depending on whether the lower of the scores from F and J is selected (labelled a, left-hand side) or the mean score rounded down of FIM® items F and J (labelled b, right-hand side) is used. In both cases, the least frequently selected FIM® category is category 1 (34% in the variant using the lower score and 30% in the variant using the mean score rounded down).

Figure 29: Transition Plots Item 8 (Toilet use) a) lower of the two scores from FIM® items F and J, b) mean score of FIM® items (rounded down if not a whole number).

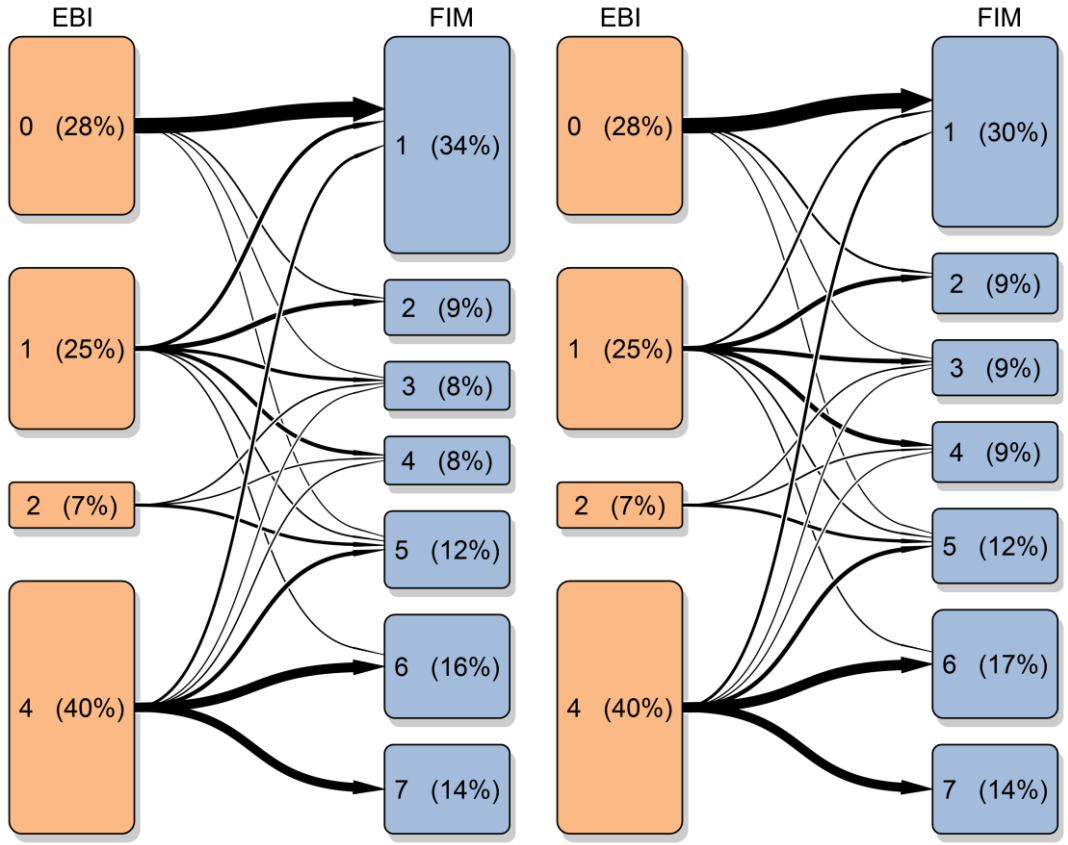


Figure 30 shows the empirical mapping of the expert variant to an ADL score in the form of a scatter plot. In the case of the ADL-from-EBI category 1, the ADL-from-FIM® scores are widely scattered, both for version a) and for b). In general, however, there is a good to very good correlation between both scores. Both regression curves are close to the black line ("perfect agreement") and the key agreement figures are good, with correlation coefficients of 0.87 and 0.88 and a weighted kappa for each of 0.87.

Figure 30: Scatter plot for item 8 from expert workshop: a) Lower score of the FIM® items F and J; b) Mean score rounded down for FIM® items F and J

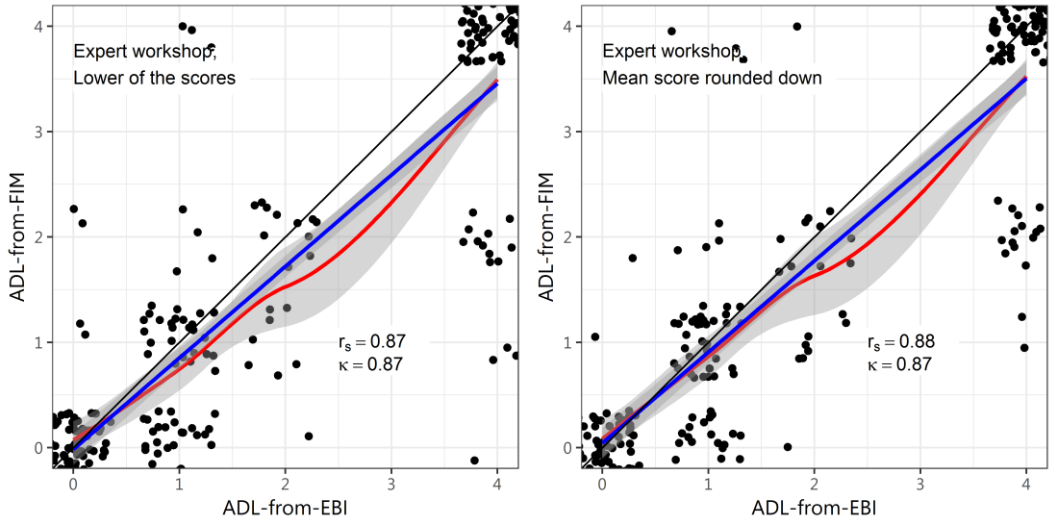
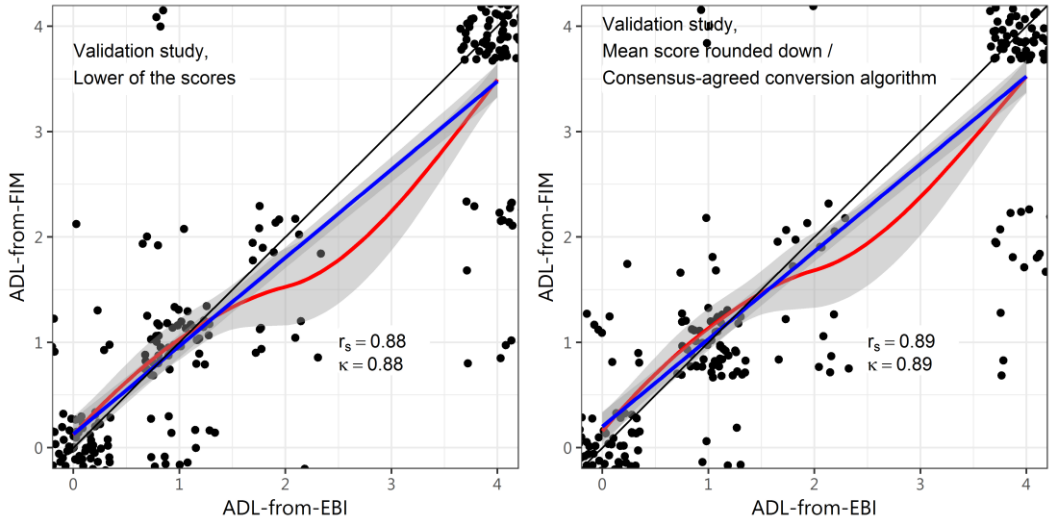


Figure 31 shows scatter plots for the mapping arising from the results of the validation study. In the case of ADL-from-EBI category 1, the ADL-from-FIM® scores are less widely scattered for both version a) and for version b). This can be put down to the change in mapping for FIM® category 2, from an ADL score of 1 to 0. In general, these versions show a close to very close correlation between the two scores. The agreement level can be viewed as very good and is somewhat better than that of the variants from the expert workshop.

Figure 31: Scatter plots for item 8 from validation study: a) Lower score of the FIM® items F and J; b) Mean score rounded down for FIM® items F and J



Item 9: Bowel management

The transition plot for item pair 9 shows that EBI response category 4 with 60% was the most frequently selected by a wide margin (Figure 32). In the case of FIM®, the two highest response categories, 6 and 7, were the most frequently selected. EBI category 3 was the least frequently selected, likewise FIM® categories 2–5, all with a maximum of 7% of patients.

The arrows indicate it is fairly clear that FIM® categories 6 and 7 should be mapped to EBI category 4. FIM® category 1 corresponds most frequently to EBI category 0. There are no clear relationships between all other FIM® categories and the EBI categories.

In the scatter plots for the two variants from the expert workshop it can be clearly seen that the ADL-from-FIM® score of 0 frequently corresponds with an ADL-from-EBI score of 2 or higher (Figure 33). In variant 1, the ADL-from-EBI score of 4 is also scattered over differing ADL-from-FIM® scores. If the proposals from the expert workshop are modified slightly, there is a better agreement between the ADL scores from FIM® and from EBI (Figure 34). The only difference in this modified version is that the lowest FIM® category is mapped to the lowest EBI category. The FIM® categories 2–4 correspond to EBI category 2. In the proposal based on the data from the validation study, the upper three FIM® categories and the upper two EBI categories all corresponding to an ADL score of 4. The level of agreement here is better than that for the proposals made at the expert workshop. The regression curves are also closer to the ideal line.

Figure 32: Transition plot for item 9 (Bowel management)

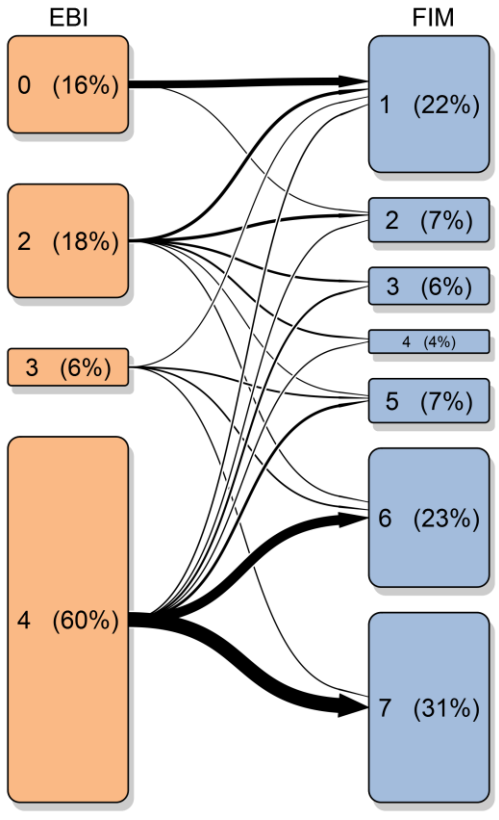


Figure 33: Scatter plot for item 9 (Bowel management) from expert workshop *variants 1 and 2*

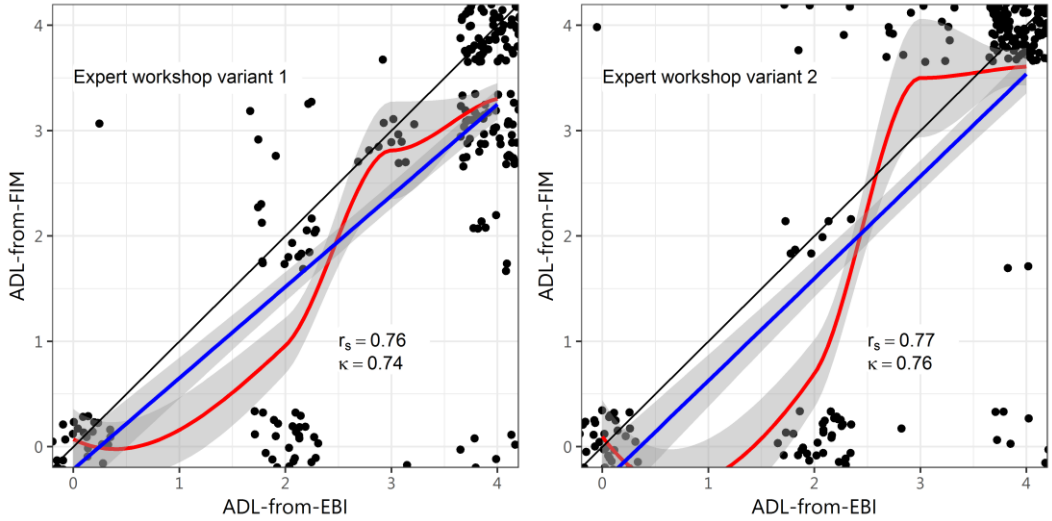
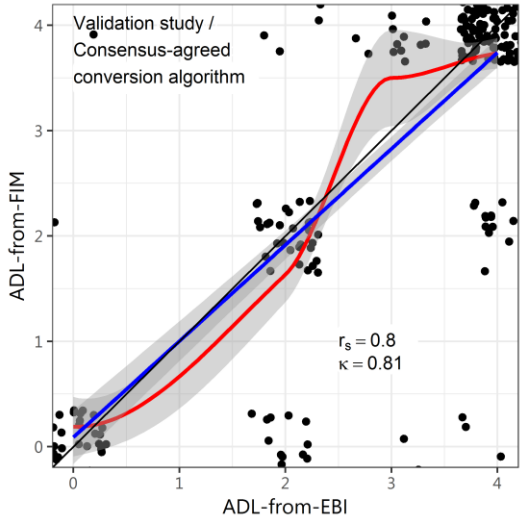


Figure 34: Scatter plot for validation study for item 9 (Bowel management)



Validation study		
EBI 9	**ADL**	**FIM® H**
0	0	1
2	2	2
		3
		4
3	4	5
4		6
		7

Item 10: Bladder management

The transition plot for item pair 9 shows that EBI response category 4 with 51% was the most frequently selected by a wide margin (Figure 35). In the case of FIM®, most frequently chosen were the lowest response category with 26% and the two highest response categories, 6 and 7, respectively with 22% and 25%. EBI category 3 with 8% was the least frequently selected, likewise FIM® categories 2–4, each with a maximum of 7% of patients.

The arrows indicate it is fairly clear that FIM® categories 6 and 7 should be mapped to EBI category 4. FIM® category 1 corresponds most frequently by a wide margin to EBI category 0. In contrast with the

proposal from the expert workshop, FIM® category 1 corresponds most frequently to EBI category 1. There are no clear relationships between all other FIM® categories and specific EBI categories.

It can be clearly seen in the scatter plot for the expert workshop that an ADL-from-EBI score of 4 corresponds to differing FIM® scores: 4 and, in particular, 3 and 1 (Figure 36). The conversion algorithm from the expert workshop does not provide for an ideal level of agreement. The lower ADL scores (0 and 1) are also not matched accurately. Two different conversion algorithms on the basis of the data from the validation study were therefore proposed, which differed slightly from those of the expert workshop.

In variant 1, deviating from the expert workshop, FIM® category 1 is mapped to EBI category 0 (Figure 37). In this new proposal, FIM® category 5 corresponds with EBI category 3. In addition, EBI categories 3 and 4 correspond with FIM® categories 6 and 7 in the ADL score conversion. This variant is otherwise unchanged from the expert proposal.

In variant 2, both of the upper EBI categories, 3 and 4, are converted to an ADL score of 4, as originally proposed at the expert workshop (Figure 38). In contrast to the proposal by the experts, however, FIM® category 5 also corresponds to an ADL score of 4. As in variant 1, FIM® category 2 is re-mapped to EBI category 1.

Both variants provide better levels of agreement in comparison with the proposal from the expert workshop, and the regression curves are closer to the ideal, with variant 2 being slightly superior.

Figure 35: Transition plot for item 10 (Bladder management)

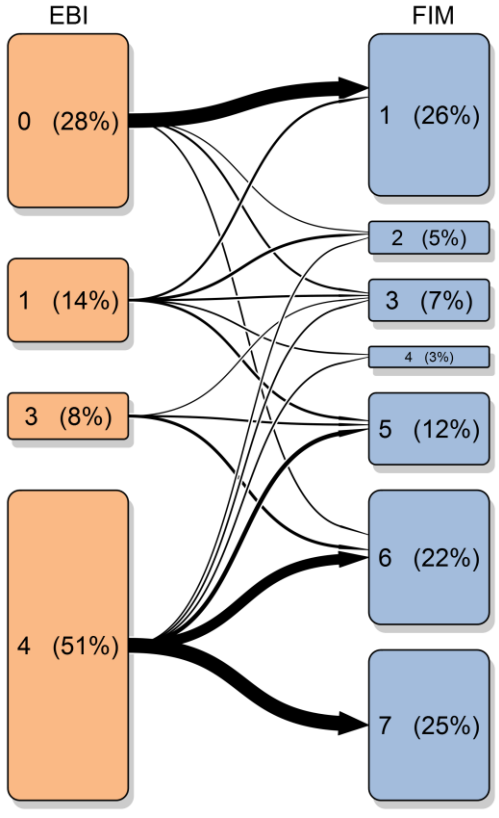
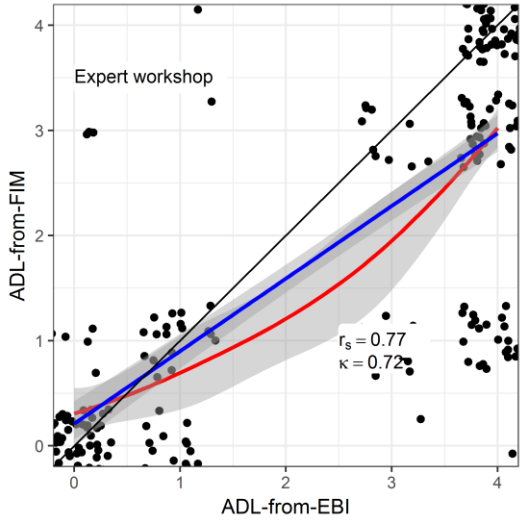
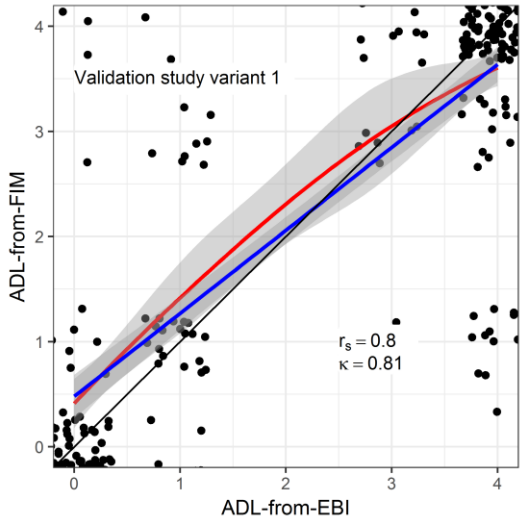


Figure 36: Scatter plot for item 10 (Bladder management) from expert workshop



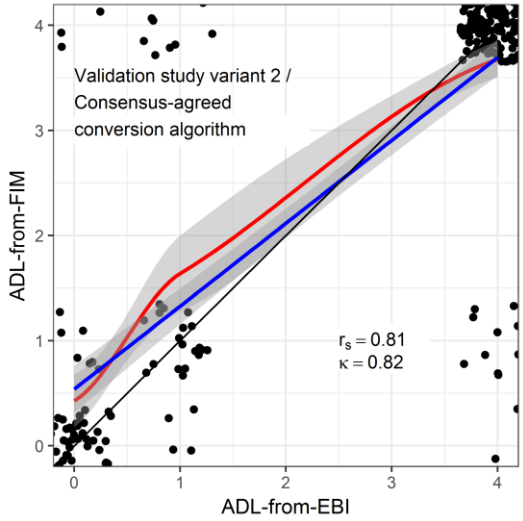
Expert workshop		
EBI 10	**ADL**	**FIM® G**
0	0	1
		2
1	1	3
		4
		5
3	3	6
4	4	7

Figure 37: Scatter plot from the validation study for *variant 1* for item 10 (Bladder management)



Validation study variant 1		
EBI 10	**ADL**	**FIM® G**
0	0	1
1	1	2
		3
		4
	3	5
3	4	6
4		7

Figure 38: Scatter plot from the validation study for *variant 2* for item 10 (Bladder management)



Validation study variant 2		
EBI 10	**ADL**	**FIM® G**
0	0	1
1	1	2
		3
		4
3	4	5
4		6
		7

Item 11: Comprehension

The transition plot for item pair 11 shows that EBI response category 4 with 48% and category 1 with 34% were most frequently selected (Figure 39). FIM® response category 7 with 34% was selected most frequently. The lowest categories were the least frequently documented, with 5% for EBI category 0 and 4% for FIM® category 1.

The arrows indicate that FIM® category 1 corresponds most frequently with EBI category 0. FIM® categories 2–4 correspond most frequently to EBI category 1. FIM® categories 6 and 7 correspond most frequently with EBI category 4. In contrast with the proposal from the expert workshop, FIM® category 1 corresponds most frequently to EBI category 1. FIM® category 5 corresponds to various EBI categories.

For both variants from the expert workshop, the scatter plot reveals a good level of agreement of the ADL-from-EBI and ADL-from-FIM® scores (Figure 40). In variant 2, the ADL-from-EBI score of 4 is, however, distributed among varying ADL-from-FIM® scores, in contrast to variant 1, in which the agreement is better (Figure 41). This can be put down to differing mappings of FIM® categories 5 and 6. The level of agreement is slightly better for variant 1 than for variant 2.

Figure 39: Transition plot for item 11 (Comprehension)

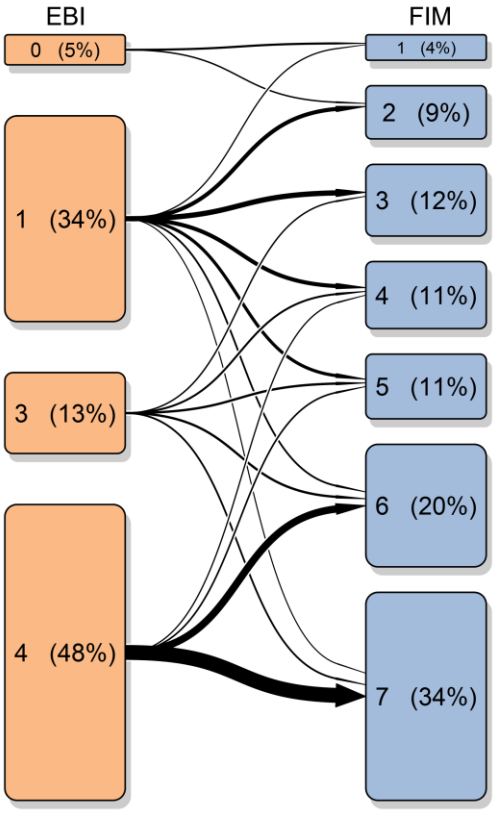
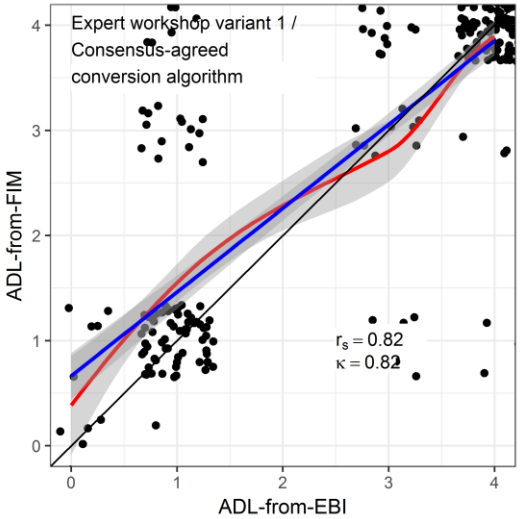
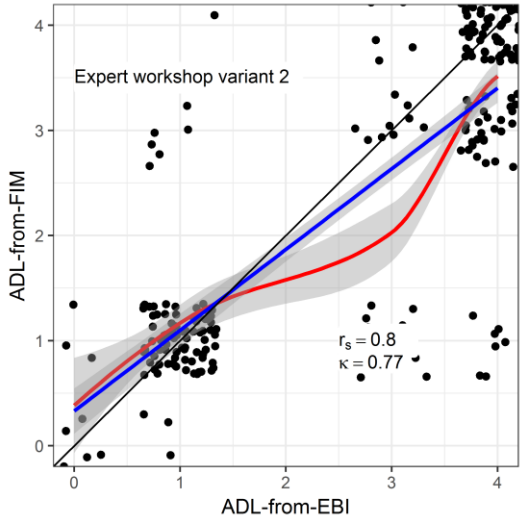


Figure 40: Scatter plot for item 11 (Comprehension) from expert workshop *variant 1*



Expert workshop variant 1		
EBI 11	**ADL**	**FIM® N**
0	0	1
1	1	2
		3
		4
3	3	5
4	4	6
		7

Figure 41: Scatter plot for item 11 (Comprehension) from expert workshop *variant 2*



Expert workshop variant 2		
EBI 11	**ADL**	**FIM® N**
0	0	1
1	1	2
		3
		4
		5
3	3	6
4	4	7

Item 12: Expression

The transition plot for item pair 12 shows that EBI response category 4 with 60% was the most frequently selected by a wide margin (Figure 42). In the case of FIM®, the two highest response categories, 6 and 7, were the most frequently documented, respectively with 20% and 33%. EBI category 3 was the least frequently selected with 5%.

The arrows indicate that FIM® category 1 corresponds most frequently with EBI categories 0 and 1. FIM® categories 2–4 correspond most frequently to EBI category 1. FIM® categories 6 and 7 correspond most frequently with EBI category 4. FIM® category 5 corresponds to various EBI categories.

For variant 2 from the expert workshop the scatter plot reveals a better agreement of the ADL-from-EBI and ADL-from-FIM® scores compared to variant 1 (Figure 43). In variant 1, the ADL-from-EBI score of 4 is distributed among varying ADL-from-FIM® scores (score of 4 as well as scores 3 and 1), in contrast to variant 2, in which the agreement is better in this area (Figure 44). However, the scores here in variant 2 are also distributed between two ADL-from-FIM® scores (4 and 1).

For this reason, a modified proposal was made on the basis of the empirical data, which is largely based on variant 2 (Figure 45). In contrast to variant 2, FIM® category 5 and the two highest FIM® categories receive an ADL score of 4. FIM® category 2 is also re-mapped to EBI category 1. In this mapping based on the validation study, the agreement parameters and the fit of the regression curves are similar to those in variant 2 from the expert workshops.

Figure 42: Transition plot for item 12 (Expression)

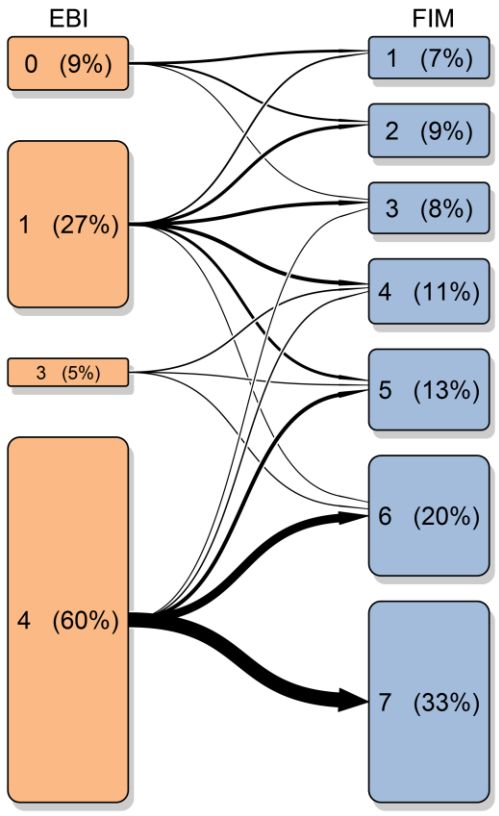
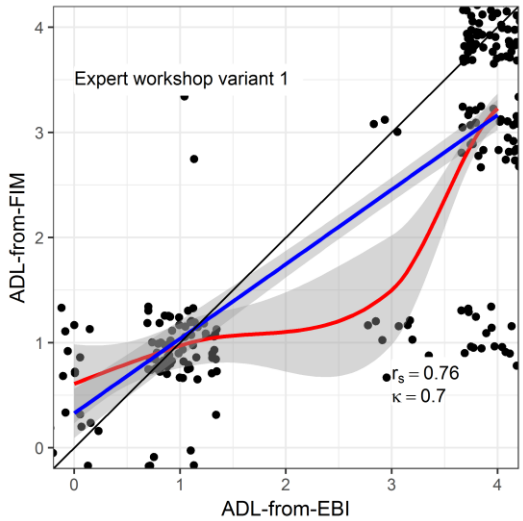
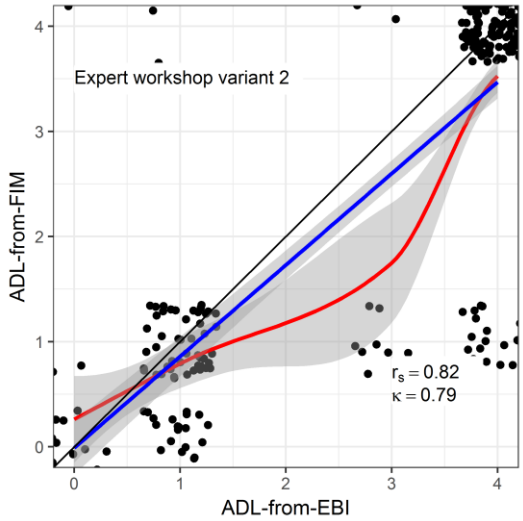


Figure 43: Scatter plot for item 12 (Expression) from expert workshop *variant 1*



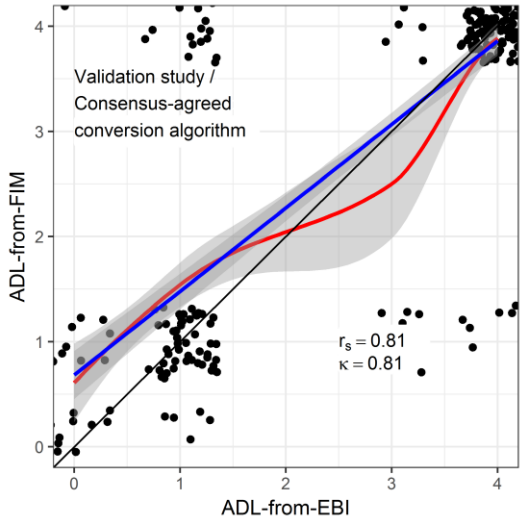
Expert workshop variant 1		
EBI 12	**ADL**	**FIM® O**
0	0	1
1	1	2
		3
		4
		5
3	3	6
4	4	7

Figure 44: Scatter plot for item 12 (Expression) from expert workshop *variant 2*



Expert workshop variant 2		
EBI 12	**ADL**	**FIM® O**
0	0	1
		2
1	1	3
		4
		5
3	4	6
4		7

Figure 45: Scatter plot for item 12 (Expression) from validation study



Validation study		
EBI 12	**ADL**	**FIM® O**
0	0	1
1	1	2
		3
		4
3	4	5
4		6
		7

Item 13: Social interaction

The transition plot for item pair 13 shows that EBI response category 4 with 66% was the most frequently selected by a wide margin (Figure 46). In the case of FIM®, the two highest response categories, 6 and 7, were the most frequently documented, respectively with 21% and 32%. The least frequent were the lowest EBI category, 0, with 5% and the lowest FIM® category with 8%.

The arrows indicate that FIM® category 1 corresponds most frequently with EBI category 0. FIM® categories 2 and 3 are most frequently mapped to EBI category 1. FIM® categories 5–7 correspond most

frequently with EBI category 4. FIM® category 4 is mapped with equal frequency to EBI categories 2 and 4.

The scatter plot for the proposal from the expert workshop shows that the two instruments are sometimes not ideally in agreement (Figure 47). It can be clearly seen, in particular, that the ADL-from-EBI score of 4 is consonant with ADL-from-FIM® scores of 4 and 2.

This led to the development of a conversion algorithm on the basis of the empirical data, which differed in two places from the expert proposal (Figure 48). In contrast to the expert proposal, both FIM® category 2 and FIM® categories 3 and 4 are given an ADL score of 2, mapping them to the middle EBI category. Also new, FIM® category 5 as well as FIM® categories 6 and 7 are mapped to EBI category 4. In this mapping based on the validation study, the levels of agreement and the fit of the regression curves are improved.

Figure 46: Transition plot for item 13 (Social interaction)

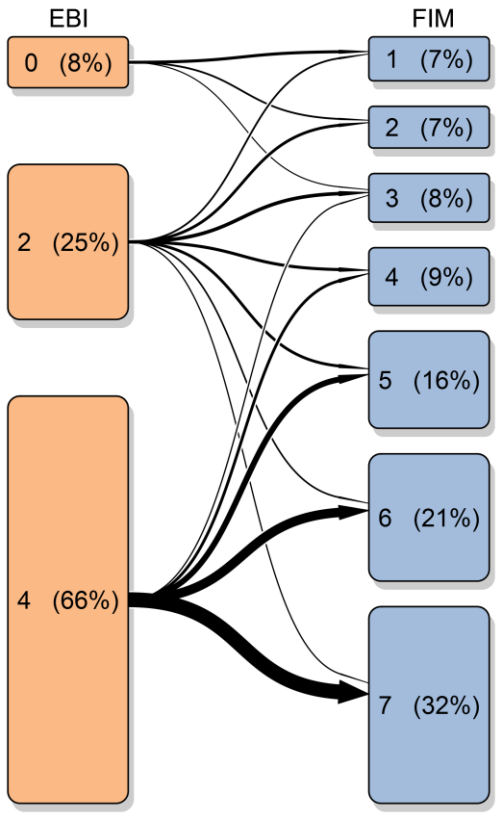
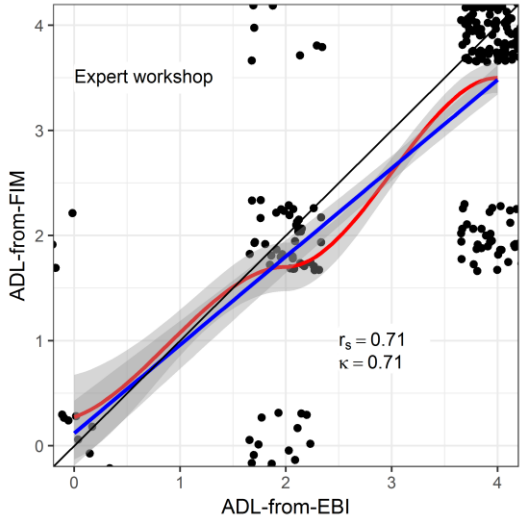
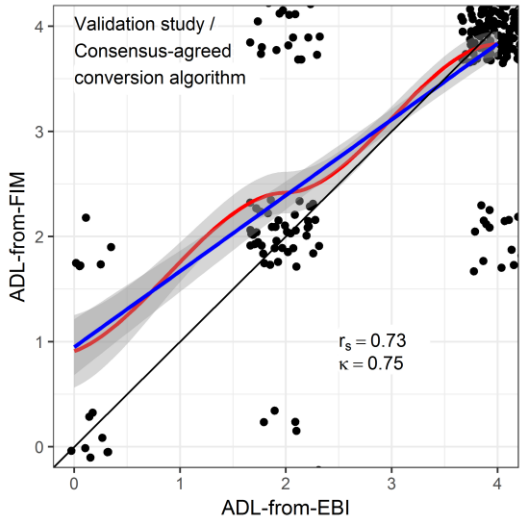


Figure 47: Scatter plot for item 13 (Social interaction) from expert workshop



Expert workshop		
EBI 13	**ADL**	**FIM® P**
0	0	1
		2
2	2	3
		4
		5
4	4	6
		7

Figure 48: Scatter plot for item 13 (Social interaction) from validation study



Validation study		
EBI 13	**ADL**	**FIM® P**
0	0	1
2	2	2
		3
		4
4	4	5
		6
		7

Item 14: Problem solving

The transition plot for item pair 14 shows that EBI response category 0 with 50% was the most frequently selected (Figure 49). The seven FIM® response categories were selected at a fairly equal frequency.

The arrows indicate that FIM® categories 1 and 2 correspond most frequently with EBI category 0. The mapping of all the FIM® categories to EBI is not particularly clear. It is not obvious from the plot, however, how the conversion algorithm from the expert workshop could be improved.

In the scatter plot the regression curves fit well to the ideal line (Figure 50). The levels of agreement are in an acceptable range.

Figure 49: Transition plot for item 14 (Problem solving)

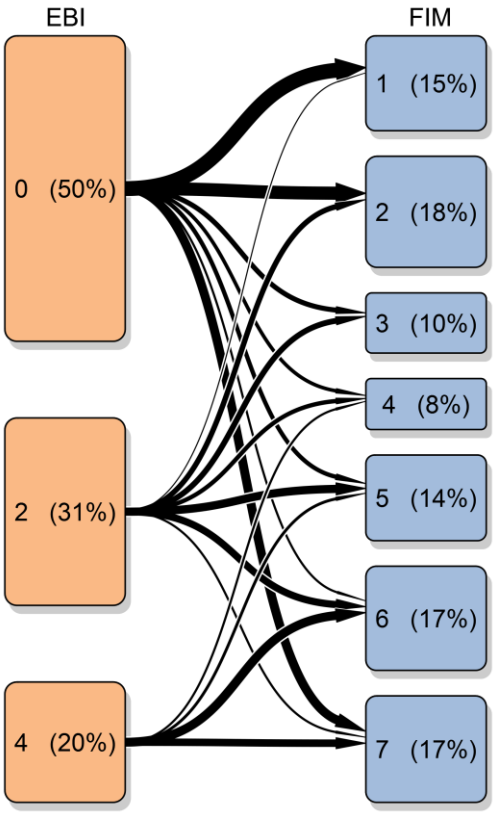
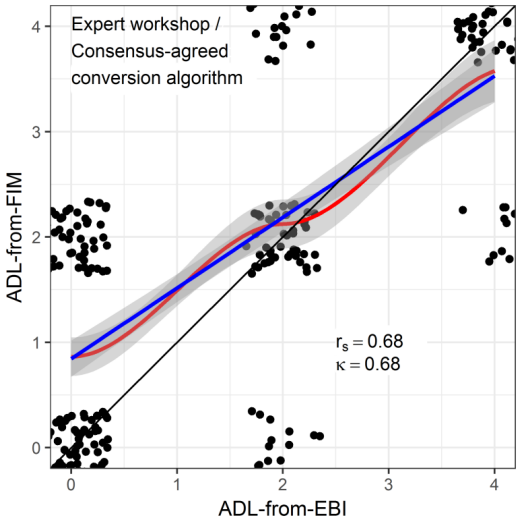


Figure 50: Scatter plot for item 14 (Problem solving) from expert workshop



Expert workshop		
EBI 14	**ADL**	**FIM® Q**
0	0	1 2
2	2	3 4 5
4	4	6 7

Item 15: Memory/learning capability/orientation

The transition plot for the last item pair, 15, shows that EBI response category 4 with 38% was the most frequently selected (Figure 51). In the case of FIM®, the two highest response categories, 6 and 7, were the most frequently documented, respectively with 18% and 24%. The lowest EBI category, 0, was the least often selected with 2%.

The arrows indicate that FIM® categories 1 and 2 correspond most frequently with EBI category 1. FIM® category 3 corresponds most frequently with EBI category 2, and FIM® category 5, with EBI category 3. FIM® category 4 is mapped with roughly equal frequency to EBI categories 2 and 3. The upper FIM® categories, 6 and 7, correspond most frequently with EBI category 4.

In the scatter plot the regression curves fit well to the ideal line (Figure 52). The level of agreement is reasonably good.

Figure 51: Transition plot for item 15 (Memory/learning capability/orientation)

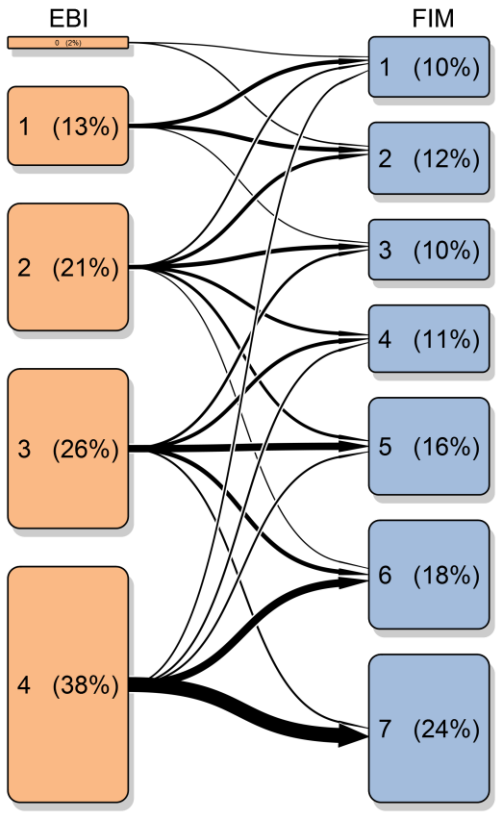
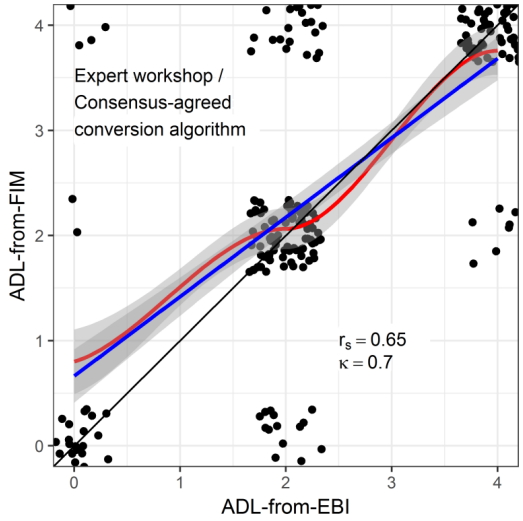


Figure 52: Scatter plot for item 15 (Memory/learning capability/orientation) from expert workshop



Expert workshop		
EBI 15	**ADL**	**FIM® R**
0	0	1
1		2
2	2	3
3		4
		5
4	4	6
		7

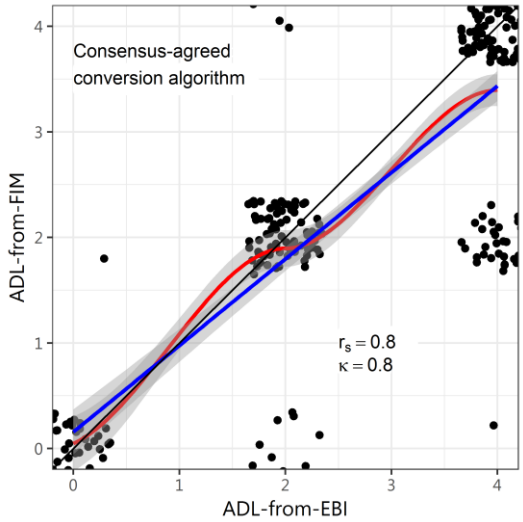
3.3. Expert consensus on the ADL score

3.3.1. Comparison at the item level

Item 1: Eating and drinking

The expert proposal modified on the basis of the validation study was judged the best fit of the variants considered. It is obvious that EBI category 3 is only rarely selected and has no clear correspondence with a specific FIM® category. FIM® cannot result in an ADL score of 3, as all of the FIM® categories correspond to one of the other EBI categories (0, 2 and 4). In numerical terms, there is almost no difference whether EBI category 3 receives an ADL score of 3 or, along with EBI category 4, an ADL score of 4. To improve the consistency with the algorithms for the other items (see items 2, 4, 9, 10 and 12) and to ensure that the scores that FIM® and EBI categories can receive are invariable (0, 2 or 4), the experts decided that EBI categories 3 and 4 should both receive an ADL score of 4. This results in almost no variation in level of agreement from that of the validation study variant. A comparison of all the conversion algorithm variants from the first expert workshop, the validation study and the second expert workshop can be found in Appendix 2.

Figure 53: Consensus-agreed scatter plot for conversion algorithm for item 1 (Eating and drinking)



Consensus-agreed conversion algorithm		
EBI 1	**ADL**	**FIM® A**
0	0	1 2
2	2	3 4 5
3 4	4	6 7

The individual key figures of the differing variants can be seen in Table 44. This shows that the consensus-agreed conversion algorithm provides the best results as confirmed by weighted kappa and the percentage agreement.

Table 44: Summary of the key figures for item 1 (Eating and drinking)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop variant 1	0,81	0,74	67,17	0,42
Expert workshop variant 2	0,81	0,76	69,43	0,47
Validation study	0,82	0,79	78,11	0,26
Consensus-agreed conversion algorithm	0,80	0,80	79,62	0,32

Item 2: Personal hygiene

The proposal to jointly give EBI categories 3 and 4 an ADL score of 4, and, in addition, to retain variant 2 from the experts ("Validation study") was accepted. The experts considered that it was most plausible that FIM® 1 and 2 should correspond to EBI category 0. An FIM® score of 2 is also only rarely given in daily clinical practice. For the experts, grouping together EBI 3 and 4 and FIM® 6 and 7 with an ADL score of 4 seemed the most reasonable suggestion for this item. The numerical differences were fairly marginal. Table 45 details the key figures for the individual variants. Even though Spearman's rank correlation coefficient and the weighted kappa value are little changed, the percentage agreement is the highest in the consensus-agreed version. In general, the level of agreement is good.

Table 45: Summary of the key figures for item 2 (Personal hygiene)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop variant 1	0,86	0,85	62,26	0,05
Expert workshop variant 2	0,87	0,86	62,26	0,22
Expert workshop variant 3	0,87	0,85	73,96	-0,14
Validation study / Consensus-agreed conversion algorithm	0,87	0,86	75,09	0,08

Item 3: Dressing and undressing

The first variant of the conversion algorithm proposed by the experts was chosen. The sub-variant using the mean score rounded down for FIM® items D and E was agreed upon as a consensus. As can be seen in Table 46, this variant provides the best empirical fit.

Table 46: Summary of the key figures for item 3 (Dressing and undressing)

	r_s : Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop variant 1 (mean score rounded down) / Consensus-agreed conversion algorithm	0,90	0,90	77,74	-0,09
Expert workshop variant 1 (lower of the scores)	0,89	0,89	76,23	0,02
Expert workshop variant 2 (mean score rounded down)	0,87	0,74	52,45	-0,47
Expert workshop variant 2 (lower of the scores)	0,87	0,78	59,25	-0,40

Item 4: Bathing/showering/taking a complete sponge bath

After intensive discussion, the conversion algorithm proposed on the basis of the validation study was chosen. The version employing the lower of the scores for FIM® items C and K was specifically selected. As can be seen in Table 47, the key figures for this version are the best, while, nevertheless, the differences in agreement as measured by the weighted kappa were, in general, minimal. The best values for percentage agreement are those from the validation study. This can be explained by the fact that ADL category 3 does not exist, and EBI responses 3 and 4, as well as FIM® categories 6 and 7, have been grouped together. The experts stressed that, with regards to content, FIM® category 5 corresponds to EBI category 2, rather than the higher EBI categories.

Table 47: Summary of the key figures for item 4 (Bathing/showering/taking a complete sponge bath)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop variant 1, lower of the scores	0,87	0,85	61,13	0,38
Expert workshop variant 1, mean score rounded down	0,85	0,85	56,60	0,17
Expert workshop variant 2, lower of the scores	0,87	0,85	61,13	0,38
Expert workshop variant 2, mean score rounded down	0,87	0,85	61,13	0,38
Validation study,lower of the score / Consensus-agreed conversion algorithm	0,88	0,88	74,72	0,23
Validation study, mean score rounded down	0,87	0,88	74,34	0,15

Item 5: Bed to chair/wheelchair transfer

The proposal from the validation study to map EBI category 0 to FIM® category 1 and 2 was rejected, as this leads neither to a plausible result with regards to content nor to a higher level of empirical agreement. As can be seen in Table 48, the original version proposed by the experts is the best variant, as measured by the weighted kappa value. For all variants, the correlations are similarly high.

Table 48: Summary of the key figures for item 5 (Bed to chair/wheelchair transfer)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop / Consensus-agreed conversion algorithm	0,85	0,85	76,98	-0,09
Validation study	0,85	0,84	78,11	-0,18

Item 6: Locomotion

The proposal from the validation study was judged the best fit of the variants considered. It can be clearly seen in Figure 26 (proposal from the validation study) on page 68 that both the regression line and the LOWESS regression curve are much closer to the black diagonal line in the case of the proposal by the experts (Figure 25). In Table 49 it can also be seen that the proposal based on the validation study generated the best correlation, weighted kappa and percentage agreement key figures.

Table 49: Summary of the key figures for item 6 (Locomotion)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop	0,81	0,77	55,85	-0,01
Validation study / Consensus-agreed conversion algorithms	0,88	0,84	63,40	0,21

Independently from the key figures, the experts also agreed that subdividing into three conversion algorithms – walking, using a wheelchair, and "both" – made particular sense for this item. This means that wheelchair users can receive a maximum of 1 ADL point, while patients who can "both" walk and use a wheelchair can receive a maximum of 2 points.

Item 7: Stairs

The original conversion algorithm proposed by the experts was considered to fit best with regards to content. The key figures from the empirical study lend weight to this decision (Table 50). A correlation of 0.82, a weighted kappa value of 0.82 and an agreement of almost 86% can be regarded as evidence of a very close agreement between the ADL scores from FIM® and from EBI. The difference in the mean ADL scores is particularly small.

Table 50: Summary of the key figures for item 7 (Stairs)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop / Consensus-agreed conversion algorithms	0,82	0,82	85,66	0,09

Item 8: Toilet use

The conversion algorithm proposed on the basis of the empirical data was agreed by the experts at consensus as the final algorithm. In this case, FIM® cateory 2 was mapped to EBI category 1. The version using the mean score rounded down was agreed upon, as this version generated the best empirical key figures.

As an exception, it should be noted that the ADL score for this item is to be specified as 0 if items 5 and 6 have also received an ADL score of 0.

Table 51: Summary of the key figures for item 8 (Toilet use)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop, lower of the scores	0,87	0,87	72,08	0,26
Expert workshop, mean score	0,88	0,87	74,72	0,21
Validation study, lower of the scores	0,88	0,88	77,74	0,17
Validation study, mean score / Consensus-agreed conversion algorithms	0,89	0,89	79,25	0,11

Item 9: Bowel management

In contrast to the variant from the expert workshop, the proposal on the basis of the validation study to give an ADL score of 4 to EBI categories 3 and 4 and to FIM® categories 5 and 7 was accepted by the experts. This also applies to the changed mapping for FIM® category 2. When this item pair was discussed, the experts once more clearly pointed out that FIM® and EBI occasionally evaluate differing aspects of rehabilitation. FIM® evaluates not only the specifics of rehabilitation but also care needs.

Table 52 shows that the key agreement values are best for the consensus-agreed conversion algorithm based on the data from the validation study. This holds true for all the parameters.

Table 52: Summary of the key figures for item 9 (Bowel management)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop variant 1	0,76	0,74	55,47	0,61
Expert workshop variant 2	0,77	0,76	71,32	0,43
Validation study / Consensus-agreed conversion algorithms	0,80	0,81	79,25	0,17

Item 10: Bladder management

Two proposals were made for this item pair based on the empirical data. After intensive discussion, the experts agreed on the second variant. The case is identical to item pair 9, "Bowel management" (see above). Thus, EBI categories 3 and 4 are grouped together and receive an ADL score of 4, as do FIM® categories 5–7. This item was considered particularly important.

Table 53 shows that the level of agreement is best for this consensus-agreed conversion algorithm when compared to the other variants. The difference in the mean ADL scores from FIM® and EBI is only -0.02 points.

Table 53: Summary of the key figures for item 10 (Bladder management)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop	0,77	0,72	58,11	0,53
Validation study variant 1	0,80	0,81	72,45	0,02
Validation study variant 2 / Consensus-agreed conversion algorithms	0,81	0,82	82,26	-0,02

Item 11: Comprehension

Two proposals had been made for this item pair at the expert workshop. The experts ended up deciding on variant 1. All the agreement parameters were better for this variant than for variant 2 (Table 54).

Table 54: Summary of the key figures for item 11 (Comprehension)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop variant 1 / Consensus-agreed conversion algorithms	0,82	0,82	75,85	-0,12
Expert workshop variant 2	0,80	0,77	66,79	0,29

Item 12: Expression

Based on the data for this item pair, a proposal modified from that of the expert workshop was submitted and subsequently agreed upon at consensus. Similar to item pairs 9 and 10, EBI categories 3 and 4 are grouped together and receive an ADL score of 4, as do FIM® categories 5–7. As FIM® and EBI are not ideally congruent with regards to content for this and for the following item pair, a broader division into three ADL categories would seem to make particular sense.

Table 55 shows that the agreement parameters are best for the consensus-agreed conversion algorithm, compared to variant 1 from the expert workshop. The differences to variant 2 are minimal according to Spearman's rank correlation coefficient and kappa. The difference in the mean ADL scores from FIM® and EBI is lowest for the consensus-agreed conversion algorithm compared to all the other variants.

Table 55: Summary of the key figures for item 12 (Expression)

	r_s : Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop variant 1	0,76	0,70	62,26	0,48
Expert workshop variant 2	0,82	0,79	76,23	0,37
Validation study / Consensus-agreed conversion algorithms	0,81	0,81	81,13	-0,11

Item 13: Social interaction

Based on the data for this item pair, a proposal modified from that of the expert workshop was submitted and subsequently agreed upon at consensus. Similar to the previous item pair, the agreement between FIM® and EBI in terms of content, as it turns out, seems to be less close than that of other items. For this item as well, the ADL score was consolidated into three levels of function. Table 56 shows that all the agreement parameters are best for the consensus-agreed conversion algorithm, compared to the proposal from the expert workshop.

Table 56: Summary of the key figures for item 13 (Social interaction)

	r_s : Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop variant 1	0,71	0,71	72,45	0,38
Validation study / Consensus-agreed conversion algorithms	0,73	0,75	81,13	-0,07

Item 14: Problem solving

The only version for this item pair is the proposal from the expert workshop. This was agreed upon by the experts at consensus. There does not seem to any way to improve this conversion algorithm. The level of agreement is reasonably good. The difference in the mean ADL scores from FIM® and EBI is minimal.

Table 57: Summary of the key figures for item 14 (Problem solving)

	r_s : Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop / Consensus-agreed conversion algorithms	0,68	0,68	61,89	-0,38

Item 15: Memory/learning capability/orientation

For the last item pair as well, the only version is the proposal from the expert workshop. This was agreed upon by the experts at consensus. The level of agreement is also moderately good. The difference in the mean ADL scores from FIM® and EBI turns out to be minimal.

Table 58: Summary of the key figures for item 15 (Memory/learning capability/orientation)

	r_s: Spearman's rank correlation	K: weighted kappa	Percentage agreement	Difference between means
Expert workshop / Consensus-agreed conversion algorithms	0,65	0,70	75,47	-0,10

3.3.2. Comparison at the overall score level

The previous sections showed how, for every item, proposals were made, deliberated on and, finally, evaluated, both empirically and with regards to content. In appendix A3 all the consensus-agreed conversion algorithms can be viewed in summary on pages 150 and 151. It is the total ADL score, however, that is of particular interest, especially its practical use. This is what is to be employed as an outcome indicator. In Figure 54, total ADL scores from EBI (x-axis) are plotted against total ADL scores from FIM® scores (y-axis). The coloured lines are regression curves showing the level of agreement between the total ADL scores from FIM® and the total ADL scores from EBI. The straight black line represents the perfect relationship, in other words the regression curve that would be obtained if ADL-from-FIM® and ADL-from-EBI were identical. As can be seen, the consensus-agreed conversion algorithm (red curve) is an improvement on the first proposal by the experts (blue curve).

Figure 54: Comparison of the developed total ADL scores

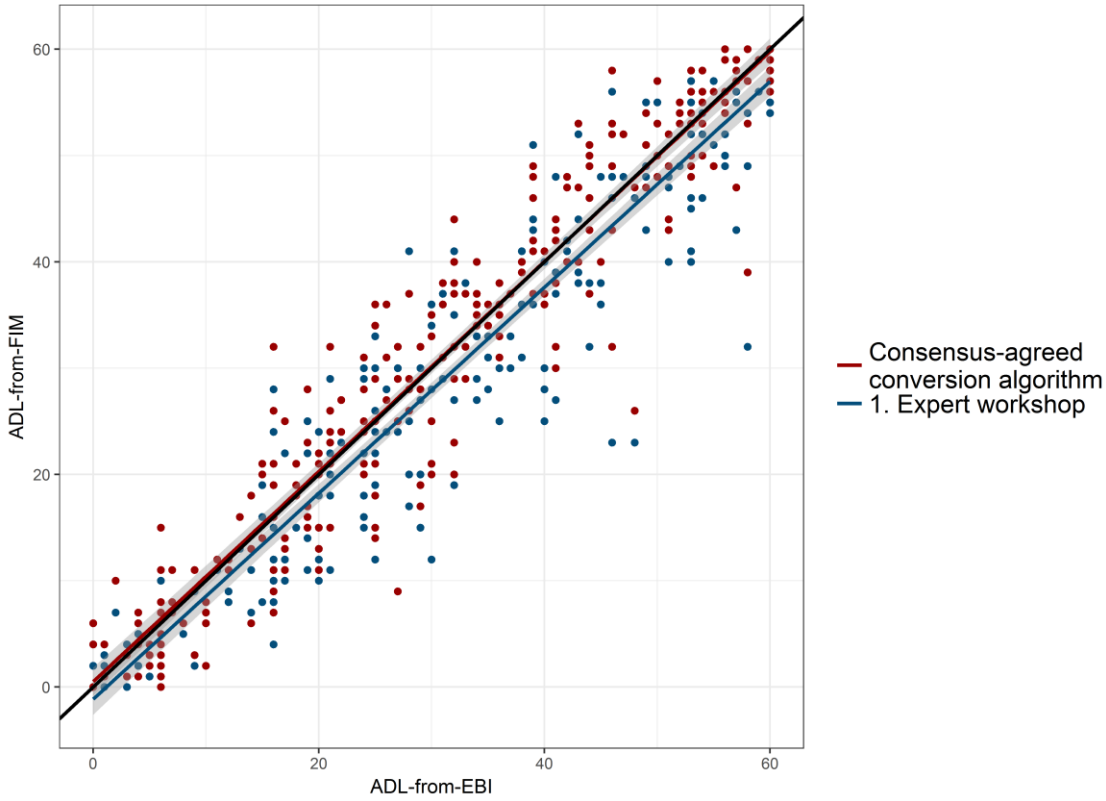


Table 59 reproduces the key figures comparing the level of agreement between ADL-from-EBI and ADL-from-FIM®. As can be seen once again, in most cases the consensus-agreed conversion algorithm generates the best results. It provides, for instance, the best Spearman rank correlation coefficient when ADL-from-FIM® is compared to ADL-from-EBI and a higher weighted kappa value.

The absolute percentage agreement between both scores is also higher for the consensus-agreed conversion algorithm (approx. 15%) than for the original proposal (approx. 11%). However, for scales with many levels of function (the range is from 0 to 60 points), the absolute percentage is not well-suited to evaluating the level of agreement. The cumulative percentage agreement within various score ranges was therefore calculated. For the variant proposed by the experts almost 37% of ADL-from-EBI and ADL-from-FIM® scores exhibited a maximal deviation of 2 points. For the consensus-agreed conversion algorithm, this figure was more than 44%. With a maximal deviation this time of +/- 4 points, the figure was 60% for the first expert variant proposed and almost 67% for the consensus-agreed version. A good 72% of the scores for the first expert variant did not deviate by more than +/- 6 points; the corresponding figure for the consensus-agreed algorithm was 81.5%.

The difference between the mean total ADL scores was 0.4 for the consensus-agreed conversion algorithm. This difference was over 2.1 points for the first algorithm proposed by the experts. A dependent (paired) sample t-test demonstrates that the probability that the mean scores in the original expert conversion genuinely differ is high ($p<0.001$). In the case of the consensus-agreed variant, on the other

hand, it is less probable that mean scores differ significantly (p=0.24). In other words, the mean ADL-from-EBI score and the mean ADL score from FIM® in the consensus-agreed version are probably equivalent.

The intraclass correlation coefficient (ICC), another measure of the consistency between two ratings, is also higher for the consensus-agreed algorithm. When linear regression is calculated, a perfect relationship is evidenced if the curve has a constant of 0 and a slope (coefficient) of 1. For the consensus-agreed version, the constant is closer to 0 (0.1 instead of -1.1), and the coefficient is closer to 1 (0.98 instead of 0.97). Furthermore, adjusted R^2, i.e. the percentage with which the ADL-from-EBI score agrees with the ADL-from-FIM® score, is somewhat higher for the consensus-agreed variant.

Table 59: Comparison of the key figures for the total ADL score

	r_s: Spearman's rank correlation r_s	K: weighted kappa	Percentage agreement (in %)	Deviation +/- 2 points (in %)	Deviation +/- 4 points (in %)	Deviation +/- 6 points (in %)
Consensus-agreed conversion algorithm	0,96	0,14	15,09	44,15	66,79	81,51
Expert workshop	0,94	0,09	10,94	36,98	60,00	72,08

	Difference between means	probability Wilcoxon-Test	ICC (unadjusted)	regression: constant	regression: coefficient	Adjustierted R^2
Consensus-agreed conversion algorithm	0,39	0,70	0,96	0,10	0,98	0,91
Expert workshop	2,14	0,00	0,94	-1,14	0,97	0,89

3.4. Plausibility check of the ADL score

The consensus-agreed ADL conversion algorithm was employed in connection with the data from the 2015 National Rehabilitation Survey. To check whether the novel ADL score is not simply a good adaptation for the comparatively small number of participants in the study, it is worth looking at the results for neurological rehabilitation from the 2015 National Rehabilitation Survey. Of particular interest is whether EBI and FIM® clinics have achieved similar results to those in the analysis when classified by the instrument they employ. One minor alteration was made to the data of the already published 2015 report in that the cases from clinic 79, in which both EBI and FIM® scores were documented, were not included in this analysis.

3.4.1. Descriptive interpretation of the outcome indicators

Both FIM® and EBI are well suited to appraising improvements in patient functional ability in neurological rehabilitation clinics. All clinics that collected FIM® (Figure 55) and/or EBI (Figure 56) scores reported, on average, an improvement between admission and discharge, though such improvement was not in every case statistically significant (an aspect that is impossible to demonstrate when confidence intervals overlap). Apparent here are also differences in ADL scores between clinics and on admission and discharge (Figure 57, EBI clinics are circled green while clinic 79, circled blue, scored both EBI and FIM® cases). All clinics registered improvements on average, even if these were not all statistically significant. Clinics whose improvements in ADL scores were not statistically significant also recorded FIM® and/or EBI improvements that were not significant. It is striking that EBI clinics (clinics 41, 64, 65, 67, 78 and 79 (clinic 79 reported both EBI and FIM® cases)) were on average neither particularly successful nor unsuccessful, but reproduce, on the whole, the results reported by the FIM® clinics. The mean total ADL score at admission was 36.84 points and at discharge 44.28 points.

Figure 55: Mean FIM® scores and the 95% confidence intervals on clinic admission and discharge (non-adjusted)

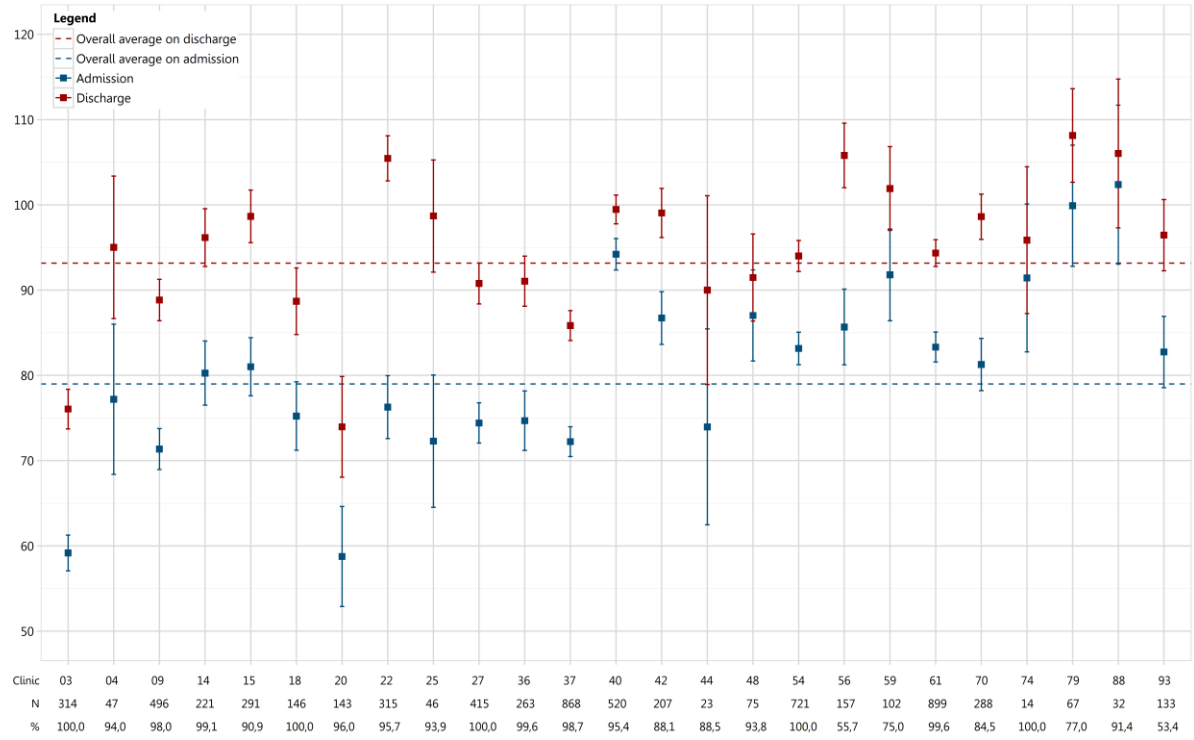


Figure 56: Mean EBI scores and the 95% confidence intervals on clinic admission and discharge (non-adjusted)

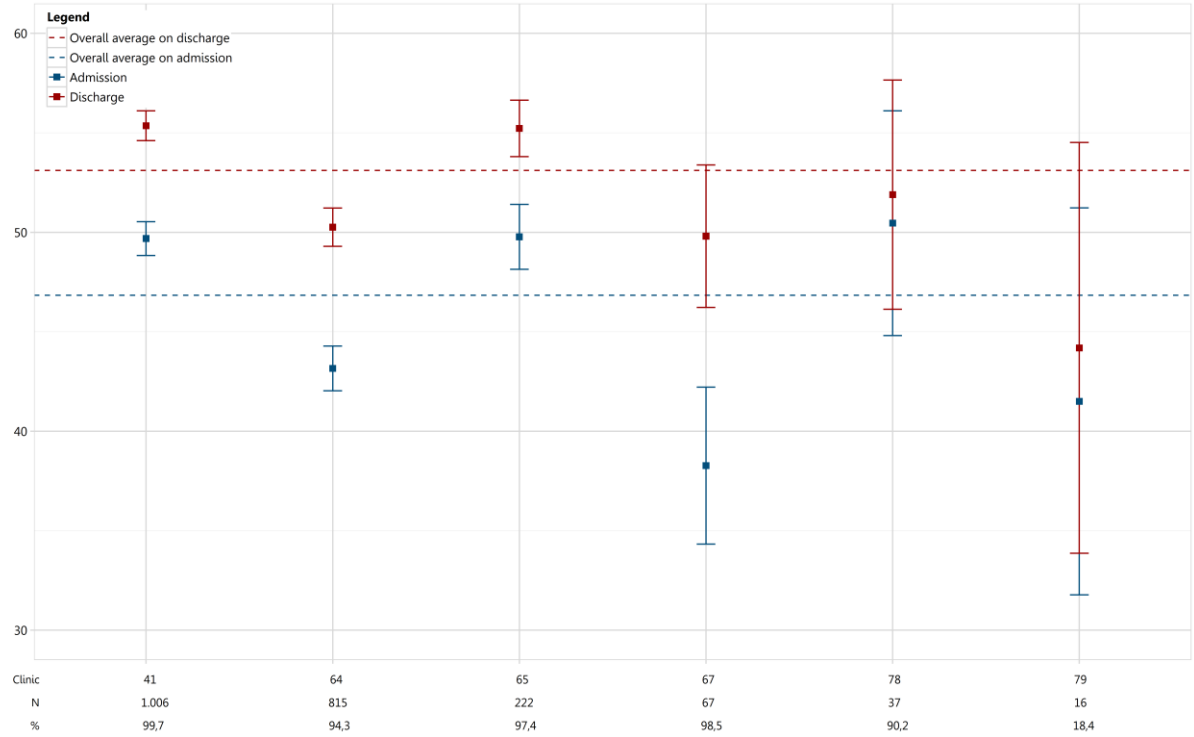
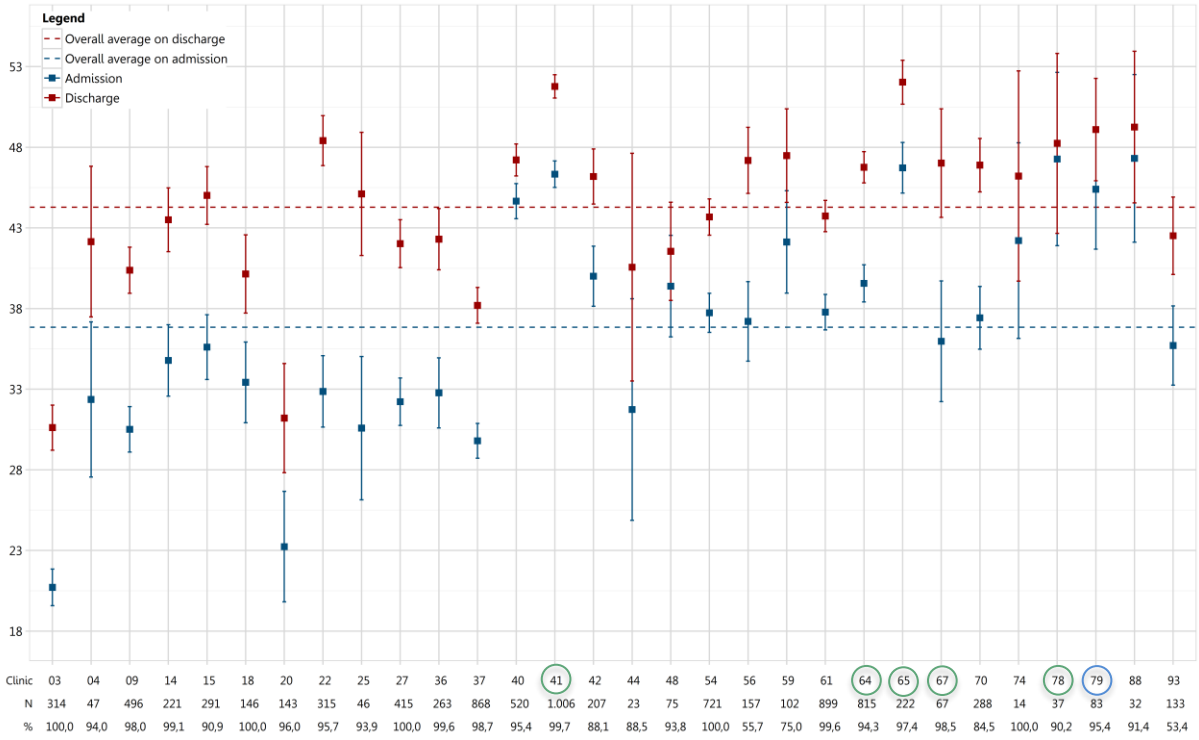


Figure 57: Mean ADL scores and the 95% confidence intervals on clinic admission and discharge (non-adjusted)



3.4.2. Risk-adjusted outcome indicators

Risk-adjusted evaluation also does not demonstrate that EBI clinics were more successful than FIM® clinics. To check for differences in the make-up of the patient cohorts in clinics, FIM®, EBI and ADL scores on discharge were analysed by means of linear regression. The explanatory variables were FIM®, EBI and ADL admission scores as well as other confounders (see Table 4). The difference between the expected scores – calculated using linear regression – and the actual scores yields so-called standardised residuals. These tend to be smaller the larger the number of cases in a clinic. Comparing FIM® clinics with each other (Figure 58) reveals that most clinics achieved results that might have been expected after taking into account the confounders. Three clinics achieved better than expected outcomes (marked as grey triangles), while four reported results that were significantly worse than expected considering the case mix (marked as grey squares). Comparing EBI clinics with each other (Figure 59) reveals that almost all clinics achieved results that might have been expected after taking into account the confounders, while one clinic had better scores than expected (marked as a grey triangle). If all clinics are compared using the novel ADL score (Figure 60) it is clear that this successfully differentiates between clinics. Most clinics achieved results that might have been expected after taking into account the confounders. Four clinics (three FIM® clinics, one EBI clinic) achieved better than expected results, while four clinics (all FIM® clinics) did worse than might be expected.

Figure 58: Funnel plot: Mean of the standardised residuals for the FIM® score on discharge by number of cases in each clinic

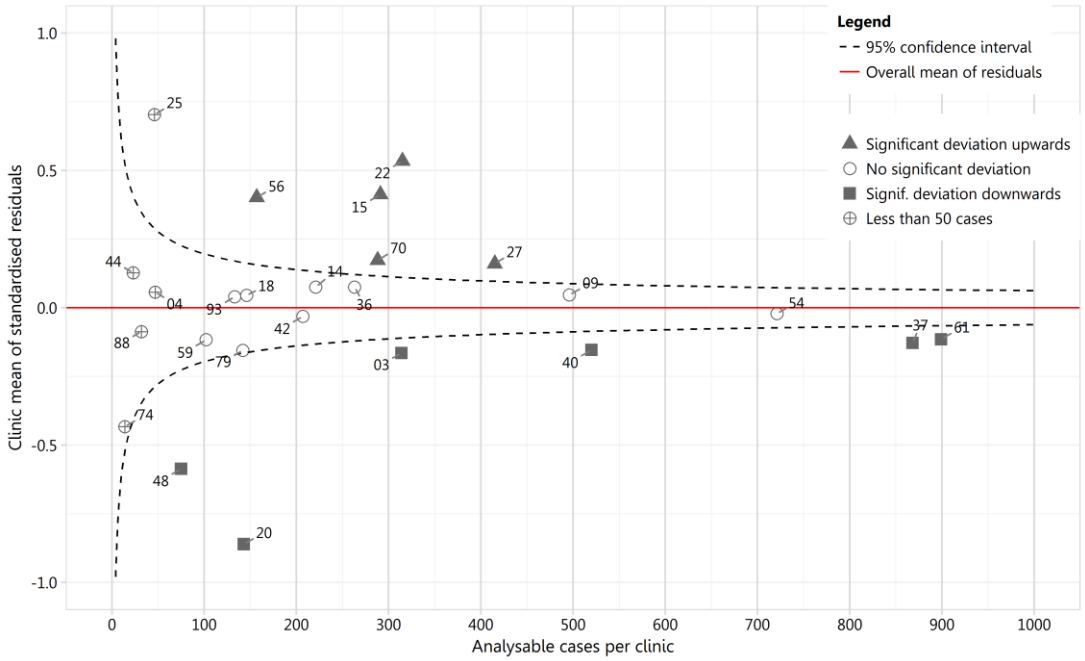


Figure 59: Funnel plot: Mean of the standardised residuals for the EBI score on discharge by number of cases in each clinic

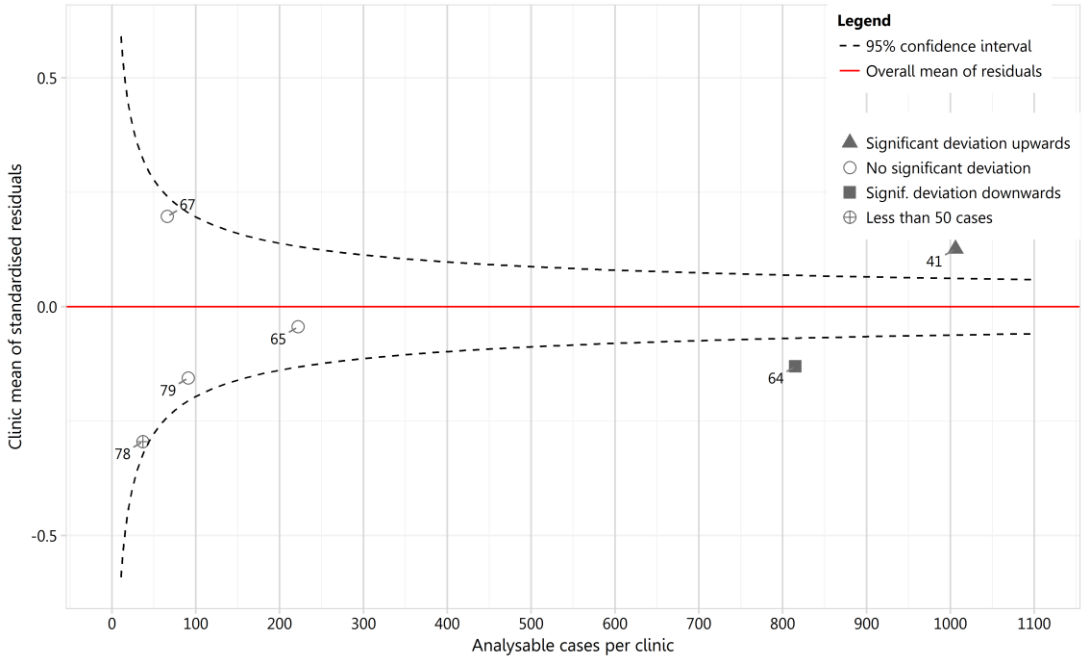
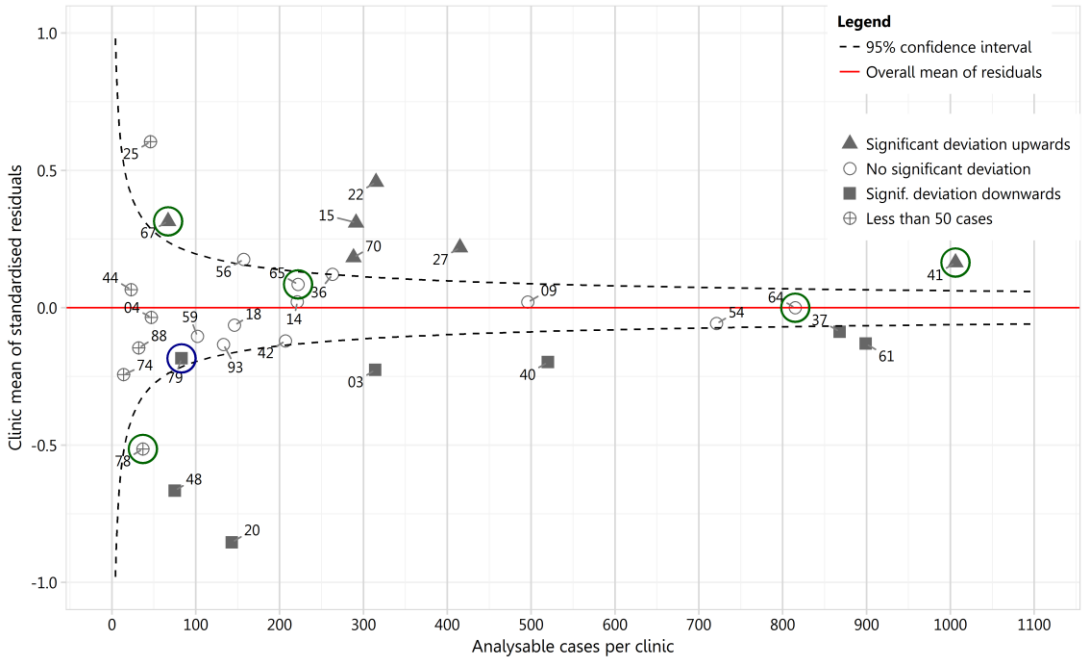


Figure 60: Funnel plot: Mean of the standardised residuals for the ADL score on discharge by number of cases in each clinic



4. Discussion

A stated aim of the National Rehabilitation Survey is to enable clinic outcomes to be compared independently of the use of FIM® or EBI within the therapeutic areas neurological, musculoskeletal and other rehabilitation programmes. This requires FIM® and EBI scores to be converted to a common scaled score. To enable the clinics to be consistently and fairly compared with regards to outcomes, the Charité was commissioned to develop a new ADL score.

Employing a mixed methods approach, the four-part study successfully developed a novel conversion algorithm that allowed FIM® and EBI scores to be converted to a common ADL score. The first step was for experts at a whole-day workshop to map analogous FIM® and EBI items and categories with regards to their content and assign to the various FIM® and EBI categories a common ADL score. At the same time, everyday functional ability of patients at five clinic sites was scored on admission to rehabilitation using both instruments to enable ADL-from-FIM® and ADL-from-EBI scores to be compared for the same individuals. Using the data collected, the conversion algorithm proposed by the experts was then validated, and small modifications were suggested based on the empirical evidence. In a further expert workshop, these modifications were discussed, and a consensus was arrived at whether they should be retained or discarded. The conversion algorithm was employed in connection with data from the 2015 National Rehabilitation Survey, demonstrating a high level of plausibility of the novel ADL score, without significant differences between clinics using either FIM® or EBI.

There is no evidence that the novel ADL score reveals a substantial difference, either positive or negative, in the quality of clinics as a result of the measuring instrument employed. Regardless of whether FIM® or EBI are employed, comparing outcomes from all clinics involved in a therapeutic area now seems feasible and meaningful.

The study has its strengths and weaknesses. As FIM® and EBI were primarily developed for patients with neurological disorders, study data was only collected at neurological rehabilitation clinics. However, FIM® and EBI are now also employed to document functional ability in areas other than neurology. It was, therefore, a requirement for the ADL score that it could also be employed in musculoskeletal and other rehabilitation programmes. The novel ADL score will be employed in the 2016 comparison reports in neurological, musculoskeletal and other rehabilitation programmes (Köhn *et al.*, 2016).

The study's multicentre approach reduces to a minimum effects due to individual collection and documentation patterns between clinics and users and enhances the universal applicability of the study's results. Clinics and study centres were selected that both employ FIM® as a standard and those that employ EBI as a standard. All study clinics were offered training in advance in the use of the instrument with which they were not familiar.

Recruiting was stratified in accordance with impairment category, ensuring that entire range of FIM® and EBI scores was covered with sufficient case numbers. This allowed FIM® and EBI scores over the whole range to be validly converted into ADL scores.

As there are valid FIM® and EBI versions in all three languages, taking into account language regions in clinic recruitment was judged, from a scientific perspective, to be of secondary importance. The study

clinics were located in individual language regions as well as in differing language regions. In particular, the study utilised German and Italian versions of FIM® and EBI, but the patients were also offered written patient information in French.

In order to replicate both instruments in the form of a common ADL score, the conversion algorithm needed to consolidate some of the recorded information. This was particularly true for FIM®, which encompasses 7 response categories for all items. ADL could only take into account a maximum of 5 ranks at the item level because to the lower number of response categories in EBI. For some items, EBI response categories also needed to be grouped together during conversion into ADL to allow consistent mapping of FIM® and EBI. Overall, a small amount of information was, therefore, also lost in the case of ADL-to-EBI conversion. Furthermore, EBI item 16 (Spatial neglect) could not be mapped to a corresponding FIM® item.

The expert discussion and the empirical data affirmed that FIM® and EBI are designed to record similar information (functional ability in activities of daily living), but that in detail they differ in orientation with regards to content. EBI focuses primarily on care and support needs in activities of daily living. FIM® records in more detail patient functional ability, with or without support from helpers or auxiliary aids. This is apparent, in particular, in terms of the number of response categories. In EBI some items only have 3 response categories, and, as a result, small improvements in functions may not always be adequately reproduced. The greater differentiation in FIM® with its 7 categories can, in turn, require more scorer experience and/or training and may reduce the validity of scores if they are not recorded strictly according to the manual.

The ADL score is based on a conversion algorithm from FIM® and EBI that allows a joint comparison of both instruments. We conclude that the study's results indicate that the novel ADL score can be employed in the future as an outcome indicator in neurological, musculoskeletal and other rehabilitation programmes. In must be borne in mind, however, that ADL is not intended to replace the original FIM® and EBI instruments. ADL is only meant to be utilised, as required, to compare these instruments. If ADL is used to perform a risk-adjusted outcome comparison, a descriptive overview of the recorded FIM® and EBI scores should also be provided.

5. References

Bortz, J., Schuster, C. (2010). Statistik für Human- und Sozialwissenschaftler. Lehrbuch mit Online-Materialien. Heidelberg, Springer.

Brünger, M., Köhn, S., Bernert, S., Wallrabe, J., Schlumbohm, A., Spyra, K. (2017). Nationaler Vergleichsbericht 2015. Neurologische Rehabilitation. Bern/Berlin, ANQ.

Bundesamt für Statistik (2016). Variablen der Medizinischen Statistik. Spezifikationen gültig ab 1.1.2016. Bern.

Charité - Universitätsmedizin Berlin (2017). Nationaler Messplan Rehabilitation (Modul 2 und Modul 3). Daten-Handbuch. Version 5.0 2017/01. Bern, ANQ.

Cleveland, W. S. (1979): Robust locally weighted regression and smoothing scatterplots. Journal of the American statistical association, 74(368). 829-836.

Cohen, J. (1968): Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychological bulletin, 70(4). 213.

Collin, C., Wade, D. T., Davies, S., Horne, V. (1988): The Barthel ADL Index: a reliability study. Int Disabil Stud, 10(2). 61-3.

Eid, M., Gollwitzer, M., Schmitt, M. (2015). Statistik und Forschungsmethoden. Weinheim, Beltz.

Gosman-Hedström, G., Svensson, E. (2000): Parallel reliability of the functional independence measure and the Barthel ADL index. Disabil Rehabil, 22(16). 702-15.

Granger, C. V., Hamilton, B. B., Keith, R. A., Zielezny, M., Sherwin, F. S. (1986): Advances in functional assessment for medical rehabilitation. Topics in Geriatric Rehabilitation, 1(3). 59-74.

Grill, E., Stucki, G., Scheuringer, M., Melvin, J. (2006): Validation of International Classification of Functioning, Disability, and Health (ICF) Core Sets for Early Postacute Rehabilitation Facilities: Comparisons with Three Other Functional Measures. American Journal of Physical Medicine & Rehabilitation, 85(8). 640-649.

Heuschmann, P. U., Kolominsky-Rabas, P. L., Nolte, C. H., Hünermund, G., Ruf, H. U., Laumeier, I., Meyrer, R., Alberti, T., Rahmann, A., Kurth, T., Berger, K. (2005): Untersuchung der Reliabilität der deutschen Version des Barthel-Index sowie Entwicklung einer postalischen und telefonischen Fassung für den Einsatz bei Schlaganfall-Patienten. Fortschr Neurol Psychiatr, 73(02). 74-82.

Hsueh, I. P., Lin, J. H., Jeng, J. S., Hsieh, C. L. (2002): Comparison of the psychometric characteristics of the functional independence measure, 5 item Barthel index, and 10 item Barthel index in patients with stroke. J Neurol Neurosurg Psychiatry, 73(2). 188-90.

Keith, R. A., Granger, C. V., Hamilton, B. B., Sherwin, F. S. (1987): The functional independence measure: a new tool for rehabilitation. Adv Clin Rehabil, 1. 6-18.

Köhn, S., Brünger, M., Schlumbohm, A., Spyra, K. (2016). Nationaler Messplan Rehabilitation (Modul 2 und 3). Auswertungskonzept. Version 1.1. Bern/Berlin, ANQ.

Kool, J., Wieser, S., Brügger, U., Dettling, M. (2009). Entwicklung eines Patientenklassifikationssystems (PCS) für die Rehabilitation in der Schweiz. Schlussbericht. Winterthur, Zürcher Hochschule für angewandte Wissenschaften.

Krol, B., Lübke, K. (2011). Wörterbuch Statistik. Die wichtigsten Begriffe mit Formeln. Dortmund, Hochschule für Oekonomie & Management.

Kwon, S., Hartzema, A. G., Duncan, P. W., Min-Lai, S. (2004): Disability measures in stroke: relationship among the Barthel Index, the Functional Independence Measure, and the Modified Rankin Scale. Stroke, 35(4). 918-23.

Landis, J. R., Koch, G. G. (1977): The measurement of observer agreement for categorical data. Biometrics, 33(1). 159-174.

Lautenschläger, J., Mau, W., Kohlmann, T., Raspe, H., Struve, F., Brückle, W., Zeidler, H. (1997): Vergleichende Evaluation einer deutschen Version des Health Assessment Questionnaires (HAQ) und des Funktionsfragebogens Hannover (FFbH). Zeitschrift für Rheumatologie, 3(56). 144-155.

Linn, B. S., Linn, M. W., Gurel, L. (1968): Cumulative illness rating scale. J Am Geriatr Soc, 16(5). 622-6.

Menzi, L. (2015): Praxisnahe Qualitätskontrolle. Schweizerische Ärztezeitung, 96(41). 1476-1477.

Menzi, L., Schwarzenbach, S., Köhn, S. (2017). Nationaler Messplan Rehabilitation (Modul 2 und Modul 3). Verfahrens-Handbuch. Version 2017/01. Bern, ANQ.
Nyein, K., McMichael, L., Turner-Stokes, L. (1999): Can a Barthel score be derived from the FIM? Clin Rehabil, 13(1). 56-63.
Prodinger, B., O'Connor, R. J., Stucki, G., Tennant, A. (2017): Establishing score equivalence of the Functional Independence Measure motor scale and the Barthel Index, utilising the International Classification of Functioning, Disability and Health and Rasch measurement theory. J Rehabil Med.
Prosiegel, M., Böttger, S., Schenk, T., König, N., Marolf, M., Vaney, C. (1996): Der Erweiterte Barthel-Index (EBI) - eine neue Skala zur Erfassung von Fähigkeitsstörungen bei neurologischen Patienten. Neurol Rehabil, 2. 7-13.
Schädler, S., Kool, J., Lüthi, H., Marks, D., Oesch, P., Pfeffer, A., Wirz, M. (2006). Assessments in der Neurorehabilitation. Bern, Huber.
Schlumbohm, A., Köhn, S., Brünger, M., Wallrabe, J., Bernert, S., Spyra, K. (2017). Nationaler Vergleichsbericht 2015. Andere Rehabilitation. Bern/Berlin, ANQ.
Sim, J., Wright, C. C. (2005): The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther, 85(3). 257-68.
Turner-Stokes, L., Williams, H., Rose, H., Harris, S., Jackson, D. (2010): Deriving a Barthel Index from the Northwick Park Dependency Scale and the Functional Independence Measure: are they equivalent? Clin Rehabil, 24(12). 1121-6.
Uniform Data System for Medical Rehabilitation (2009). The FIM System Clinical Guide, Version 5.2. Buffalo, UDSMR.
Vanbellingen, T., Ottiger, B., Pflugshaupt, T., Mehrholz, J., Bohlhalter, S., Nef, T., Nyffeler, T. (2016): The Responsiveness of the Lucerne ICF-Based Multidisciplinary Observation Scale: A Comparison with the Functional Independence Measure and the Barthel Index. Front Neurol, 7. 152.
Wade, D. T. (1992). Measurement in neurological rehabilitation. Oxford, Oxford University Press.
White, D. K., Wilson, J. C., Keysor, J. J. (2011): Measures of adult general functional status: SF-36 Physical Functioning Subscale (PF-10), Health Assessment Questionnaire (HAQ), Modified Health Assessment Questionnaire (MHAQ), Katz Index of Independence in activities of daily living, Functional Independence Measure (FIM), and Osteoarthritis-Function-Computer Adaptive Test (OA-Function-CAT). Arthritis Care Res (Hoboken), 63 Suppl 11. S297-307.

Acknowledgements

It would not have been possible to undertake this study without the active support of many people. We would like to thank all the clinic personnel involved in the study. As representatives of everyone who took part, we would like to thank by name the local study coordinators Myrta Kohler (Valens Clinics), Petra Wildisen (Clinica Hildebrand, Brissago), Britta Lassen and Andrea Sandrock (RehaClinic Bad Zurzach and Kilchberg), and Ruth Plüss and Claudia Pracht (Klinik Bethesda, Tschugg). Many thanks, as well, to the experts involved in both workshops for the stimulating and constructive discussion.

In addition, our thanks to Luise Menzi and Stephanie Studer from the ANQ office for helping organise the study, in particular in recruiting the study clinics and experts, and in managing the expert workshops. Finally, we would like to thank the student assistants from the Charité – Universitätsmedizin Berlin for, in particular, their support in data preparation and reporting.

Glossary

Dependent variable: A parameter that can vary with changes to other non-related variables, e.g. age, comorbidity. When comparing clinics, the dependent variable is the selected outcome indicator, for instance, EBI.

Adjustment: See risk adjustment.

ADL score: The ADL score measures functional ability in important Activities of Daily Living (ADL). The scores are generated by means of a conversion algorithm from FIM® and EBI, with the aim of comparing outcomes from all rehabilitation clinics in the areas of neurological, musculoskeletal and other rehabilitation programmes independently of the instrument employed.

Bar chart: Chart graphically presenting frequencies of parameters by means of horizontal bars. Frequencies of different individual parameters can also be stacked in a vertical bar (stacked bar chart). See also, column chart.

Box plot: A method of representing numerical data in graph form, such as age in years, to give a rapid impression of how the data is distributed. Each box shows the median, arithmetic mean, 25th percentile and 75th percentile. A box represents the mid 50% of the data. Data that lies outside the box (25% above and 25% below) is represented by "whiskers", while points represent outliers.

Case mix: Composition of the patient population, e.g. sociodemographic factors, comorbidity, diagnoses).

Confounders: Confounders are factors that can influence both dependent and independent variables, e.g. age, comorbidity. Their effect can be controlled by means of statistical risk adjustment.

Cumulative Illness Rating Scale (CIRS): CIRS is a third-party instrument used to record comorbidity (Linn *et al.*, 1968). Each of the instrument's 14 organ systems is scored on a 5 point scale from 0 ("No problem") through to 4 ("Extremely severe") by medical personnel. The total CIRS score is in a range of 0 points ("No comorbidity") through to 56 points ("Maximum possible comorbidity").

Expected value: A value estimated by means of regression analysis of the case mix (in other words, the independent variables).

Extended Barthel Index (EBI): Like FIM®, EBI records functional ability in important activities of life and was developed as an alternative to FIM® (Prosiegel *et al.*, 1996). Per item, 0 ("Unable") to 4 ("Independent") points can be given, and the total score from the individual items is therefore in a range from 0 (maximum impairment) to 64 (minimal impairment) points. EBI can be employed as an outcome indicator in neurological, musculoskeletal and other rehabilitation programmes. ADL scores calculated from EBI (and from FIM®) can be used to perform a risk-adjusted comparison of outcomes from clinics.

Case: A patient discharged during the period of data collection (calendar year).

Number of cases: Number of cases used as a basis for analysis or data characterisation.

Error bar chart: Representation in graph form of numerical data used, for instance, to visualise mean scores with confidence intervals.

FIM® instrument/Functional Independence Measure (FIM®): FIM® records, with its 18 items, functional ability in important Activities of Daily Living (ADL) on a seven point response scale, from 1 ("Total assistance") to 7 ("Complete independence") (Keith et al., 1987). Adding scores from all the items results in a total score ranging from 18 (maximal impairment) to 126 points (minimal impairment). FIM® can be employed as an outcome indicator in neurological, musculoskeletal and other rehabilitation programmes. ADL scores calculated from FIM® (and from EBI) can be used to perform a risk-adjusted comparison of outcomes from clinics.

Funnel plot: Representation in graph form of numerical data plotted against the number of cases. Similar to error bar charts, in this report the funnel plots show standardised residuals, which on the y-axis[x-axis?] are plotted against the number of cases in the analysis for each individual clinic. In this way, all possible relationships between outcomes and clinic sizes can be taken into account.

Statistical population: Total number of cases.

Health Assessment Questionnaire (HAQ): HAQ records physical impairment in activities of daily living, and is employed as an outcome indicator in musculoskeletal rehabilitation (Lautenschläger *et al.*, 1997; White *et al.*, 2011). In total, it comprises 24 items, each scored from 0 points (best quality of life) to 3 points (worst quality of life).

Histogram: A histogram represents in graph form the frequency distribution of numerical data, e.g. age in years. The areas of the rectangles are proportional to the frequency of each of the parameter classes (e.g. 1 year classes for age).

Item: A single question in a questionnaire or assessment form, e.g. in EBI.

Intraclass correlation coefficient (ICC): ICC, similar to Cohen's kappa, is a measure of the correlation between two evaluations. It is particularly well suited to comparing interval-scaled data, e.g. ADL scores from FIM® and from EBI.

Kappa: Cohen's kappa is a statistical instrument which measures the correlation of two evaluations, for instance made by two people or two instruments, and, generally, has a value of 0 to 1 [-1 to 1?]. A value of 0 indicates that the agreement is purely random. A value of 1 indicates an exact agreement in all cases.

Confidence interval (CI): The CI describes the precision of the estimated value of a parameter, e.g. the mean value. In the case of a 95% confidence interval, the probability that the range of values contains the correct value is 95%.

Spearman's rank correlation: Spearman's rank correlation is a measure of agreement between two variables. In this study, the agreement between the calculated ADL scores from FIM® and the ADL scores from EBI is analysed. A score of 0 indicates no linear relationship, while +1 is a perfect match.

Maximum: The highest recorded score.

Median: Score lying at the midpoint of a distribution of metric data, e.g. age. Half of the evaluated scores lie above and below the median (equivalent to the 50th percentile).

Baseline medical statistical data: Portion of the data collected by the Swiss Federal Statistical Office (SFSO) containing socio-demographic data such as details of hospitalisation, and inpatient diagnosis

and treatment costs. Data from this source, such as primary diagnoses and age, are used to adjust the clinic comparison to take into account the case mix.

Minimum: The lowest recorded score.

Mean score: The arithmetic mean (average) of recorded scores.

Outcome: End result of a rehabilitation programme as measured by outcome indicators such as EBI.

Percentile: A percentile indicates the percentage of observed variables that are below a particular value. At the 25th percentile 25% of observations are below this value, at the 75th percentile 75% of observations are below.

Predictor: A variable used to predict another value. See also confounders and independent variables.

Regression: Statistical procedure used to estimate a dependent variable (outcome) on the basis of one or more independent variables (predictors). In this report the dependent variable is estimated using *linear* regression, as the relationship between the variables is expected to be linear.

Residual (residual score): Difference between the score as estimated using predictors and the actual recorded score in a treatment case. The effect confounders have on residuals can be eliminated, thus removing distortions due to variations in the composition of patient populations at the clinics involved. Above-average residual values indicate high quality, indicating the outcome is better than might be expected given the case mix.

Risk adjustment: Statistical method of adjusting for the effects of those confounders that clinics cannot control. Once the calculated parameters have been appropriately adjusted, clinics can be more fairly compared. The principal confounders are those that arise from the varied composition of the case mix.

Column chart: Chart that represents in graph form the frequencies of parameters by means of vertical bars. See also, bar chart.

Significance: Differences between measured variables are described as significant if the probability they have arisen by chance is not above a specific defined threshold. The maximal permitted probability of error is termed the significance level or α.

Standard deviation (SD): The range of dispersal of a set of values of a variable around their mean. The standard deviation is equal to the square root of the variance, and, along with the mean and number of cases, is required to calculate the confidence interval.

Standardised residual: Residuals that have been standardised to ensure the standard deviation of the residuals is 1 and the mean is 0.

Sample population: Subset of a whole given population. Statistical techniques can be used to extrapolate the results for a sample population for the population as a whole.

Scatter plot: A scatter plot is a representation in graph form of the measured values of two variables. In this case, the ADL score from FIM® is plotted on the y-axis while the ADL score from EBI is on the x-axis.

Actual value (measured value): The actual measured value. Often compared with the expected value. The difference between the two is termed the residual.

Transition plot: Transition plots are representations in graph form of the distribution of responses to two variables, in this case the corresponding FIM® and EBI items. Each FIM® and EBI response category

is represented by a rectangular column. The length of the column indicates the frequency of the particular response category in the sample population. The FIM® and EBI columns are linked by arrows, indicating which EBI category/ies were selected in the presence of a given FIM® category. The thickness of the arrows symbolises the frequency of each connection. The more frequently an EBI category was selected in combination with a particular FIM® category, the thicker the arrow.

t-test for dependent samples: Statistical test used to assess the null hypothesis that there is no difference between two expected values of the underlying distributions in a population.

Independent variable: A parameter that can be influenced by independent variables. When measuring outcomes, the independent variable is also termed the predictor.

Variable: Statistical parameter (e.g. place of residence before admission), which assigns levels (e.g. "intensive care unit" or "at home") to parameter holders (patients).

Variance: An indicator of the range of distribution of measured values. It is calculated as the squared deviation of the individual values from their mean. The standard deviation is the square root of the variance.

The definitions of the above terms have been formulated using language that is easily understandable by a wide range of readers. They may be simplified and may not completely mirror in every detail the definitions based on current statistical standards. Please refer to the literature for more detailed definitions of statistical terms (Bortz, Schuster, 2010; Krol, Lübke, 2011; Eid *et al.*, 2015).

List of figures

List of tables

List of abbreviations

ADL	Activities of Daily Living
ANQ	Swiss National Association for Quality Development in Clinics and Hospitals
SFSO	Swiss Federal Statistical Office
CIRS	Cumulative Illness Rating Scale (measure of comorbidity)
EBI	Extended Barthel Index
CID	Case identification number
FIM®	FIM® Instrument/Functional Independence Measure
HAQ	Health Assessment Questionnaire
ICC	Intraclass correlation coefficient
ICD-10	International Statistical Classification of Diseases and Related Health Problems, 10th Revision
ICF	International Classification of Functioning, Disability and Health
κ	Kappa
KEK	Ethics Commission of the Canton of Bern
LOWESS	Locally Weighted Scatterplot Smoothing (modelling for regression curves)
n	Number of cases
r_s	Spearman's rank correlation

Appendix

A1 Sample population description - clinic compariso

Figure 61: Age by clinic

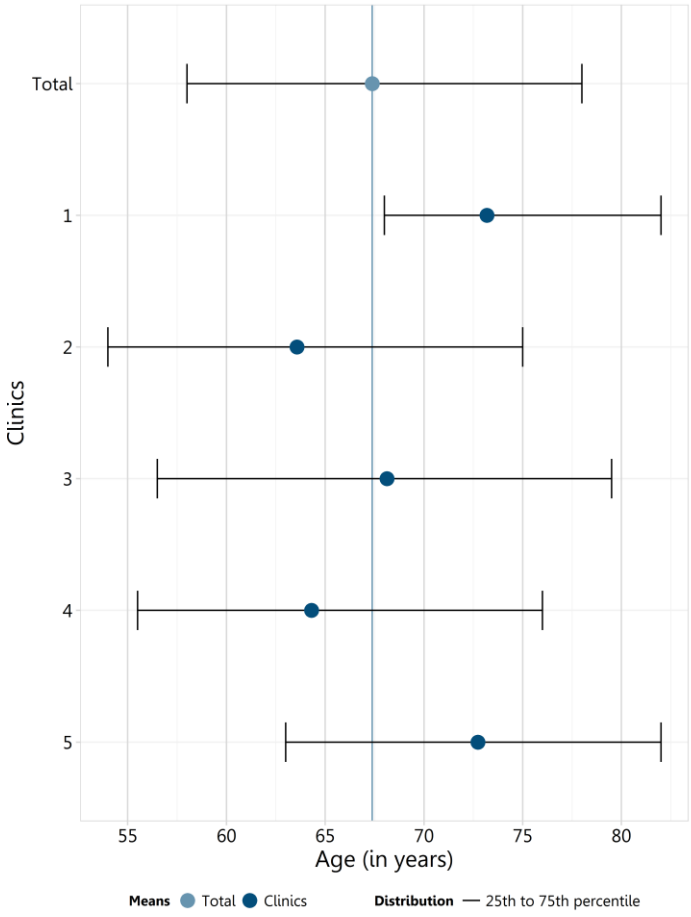


Figure 62: Gender by clinic

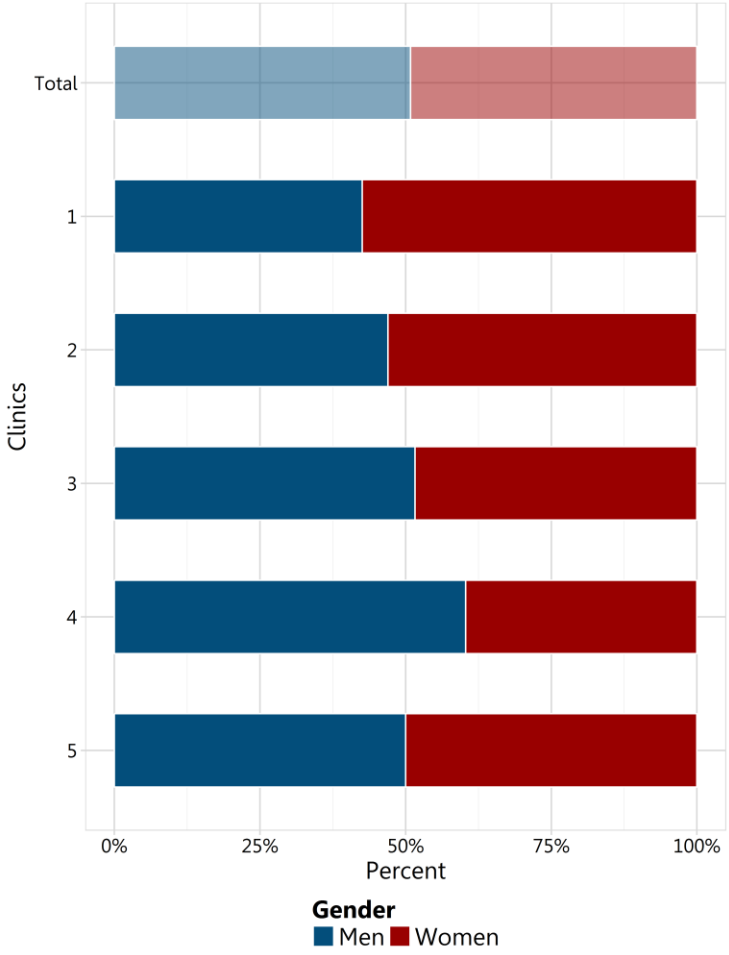


Figure 63: Nationality by clinic

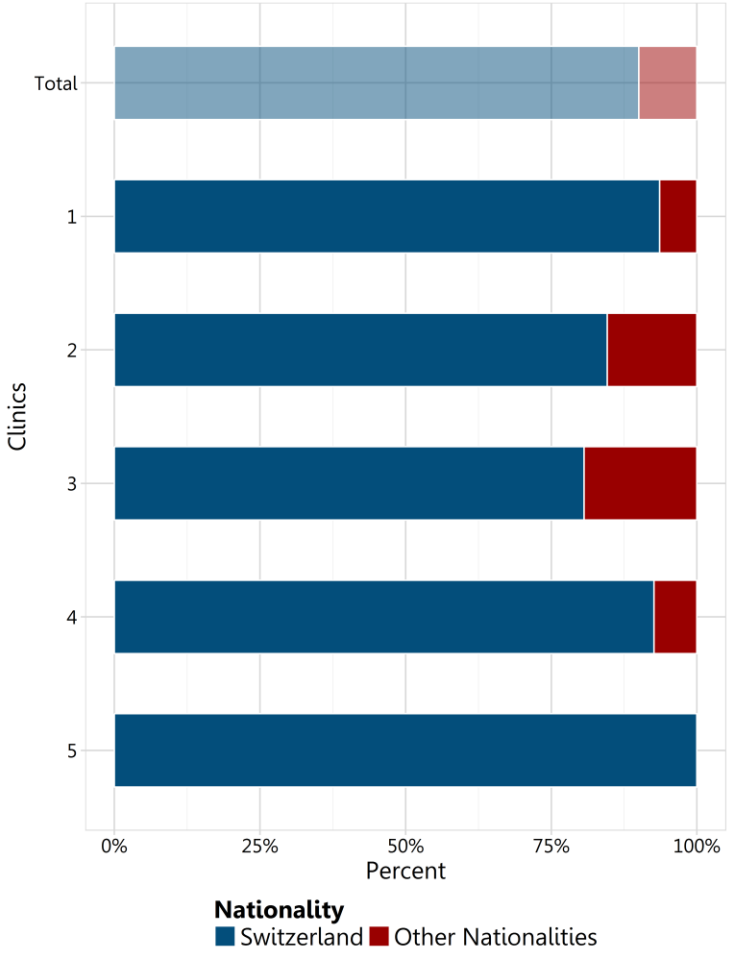


Figure 64: Comorbidity (CIRS) by clinic

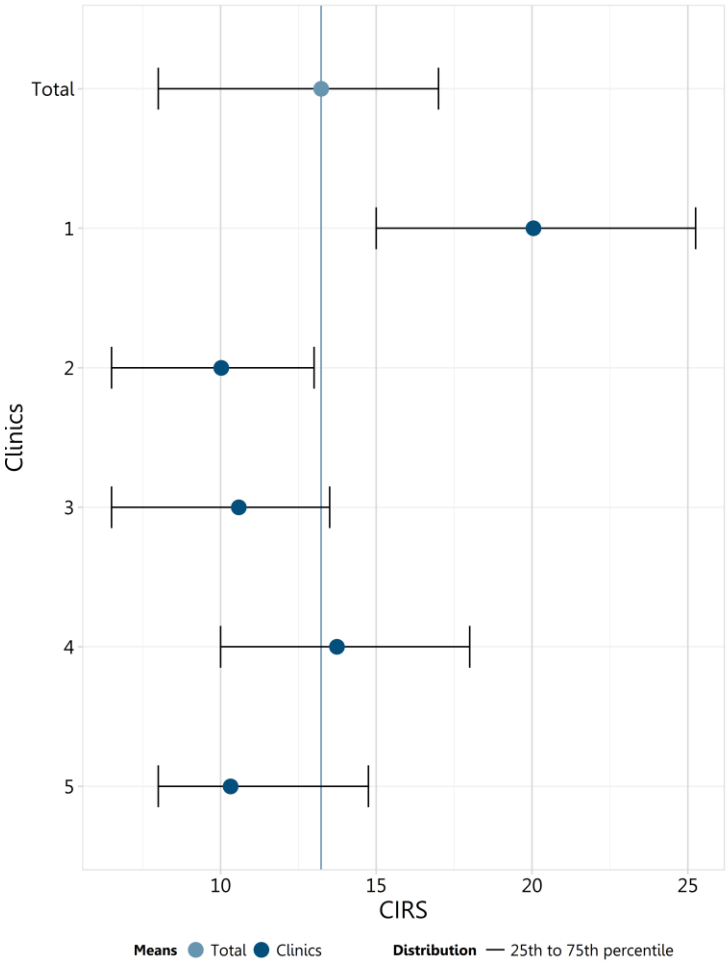


Figure 65: EBI categories by clinic

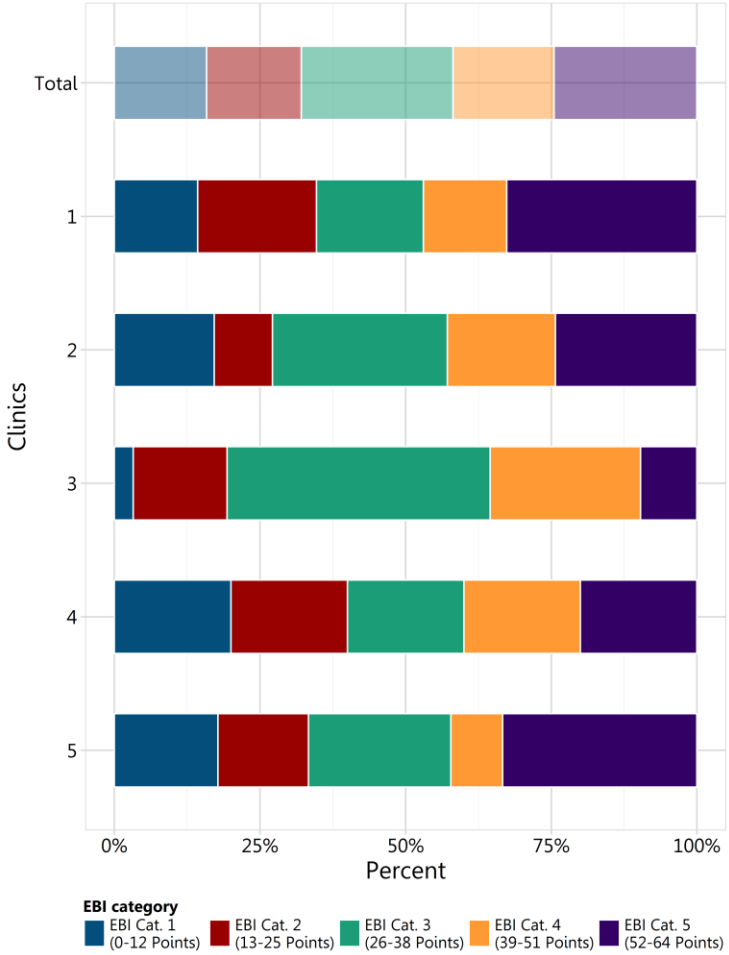
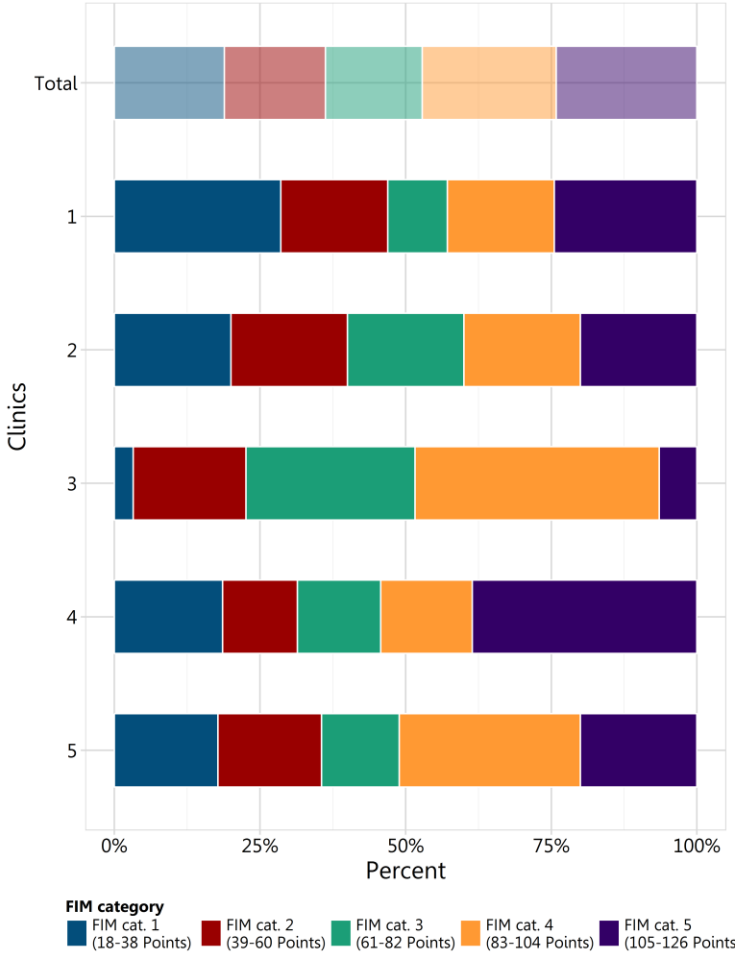


Figure 66: FIM® categories by clinic



A2 Comparison of scatter plots and conversion algorithms for all variants

This section contains a comparison of scatter plots and key agreement parameters for all variants of item mappings. Shown are the results for the conversion algorithms from the first expert workshop (sometimes in several variants), the results on the basis of the validation study and the results of the consensus-agreed conversion algorithms from the second expert workshop. The results from each conversion algorithm are appropriately labelled.

Item 1: Eating and drinking

Figure 67: Scatter plot item 1 (Eating and drinking)

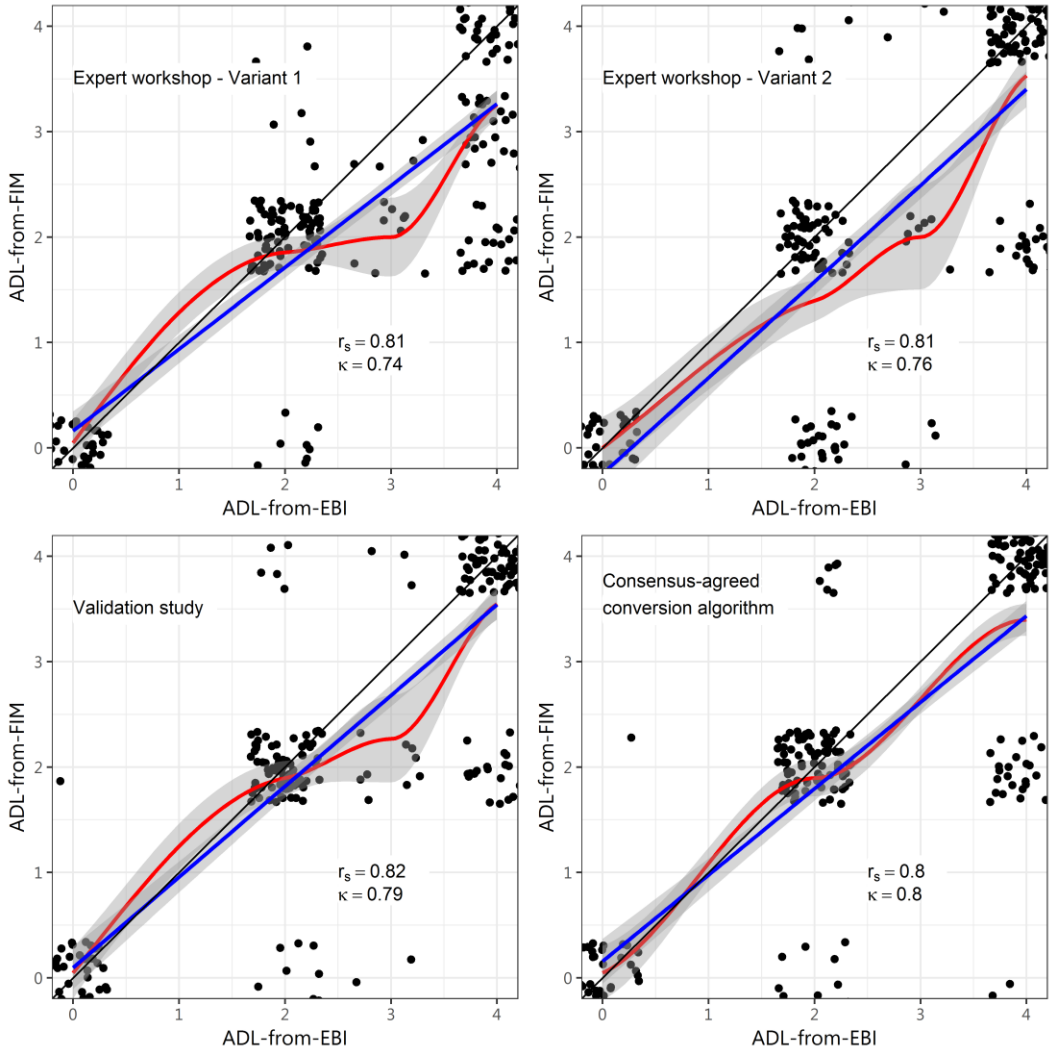


Table 60: Overview of conversion algorithm for item

Expert workshop variant 1		
EBI 1	**ADL**	**FIM® A**
0	0	1
		2
2	2	3
		4
		5
3	3	6
4	4	7

Expert workshop variant 2		
EBI 1	**ADL**	**FIM® A**
0	0	1
		2
		3
		4
2	2	5
3	3	
4	4	6, 7

Validation study		
EBI 1	**ADL**	**FIM® A**
0	0	1
		2
2	2	3
		4
		5
3	3	
4	4	6,7

Consensus-agreed conversion algorithm		
EBI 1	**ADL**	**FIM® A**
0	0	1
		2
2	2	3
		4
		5
3	4	6
4		7

Item 2: Personal hygiene

Figure 68: Scatter plot item 2 (Personal hygiene)

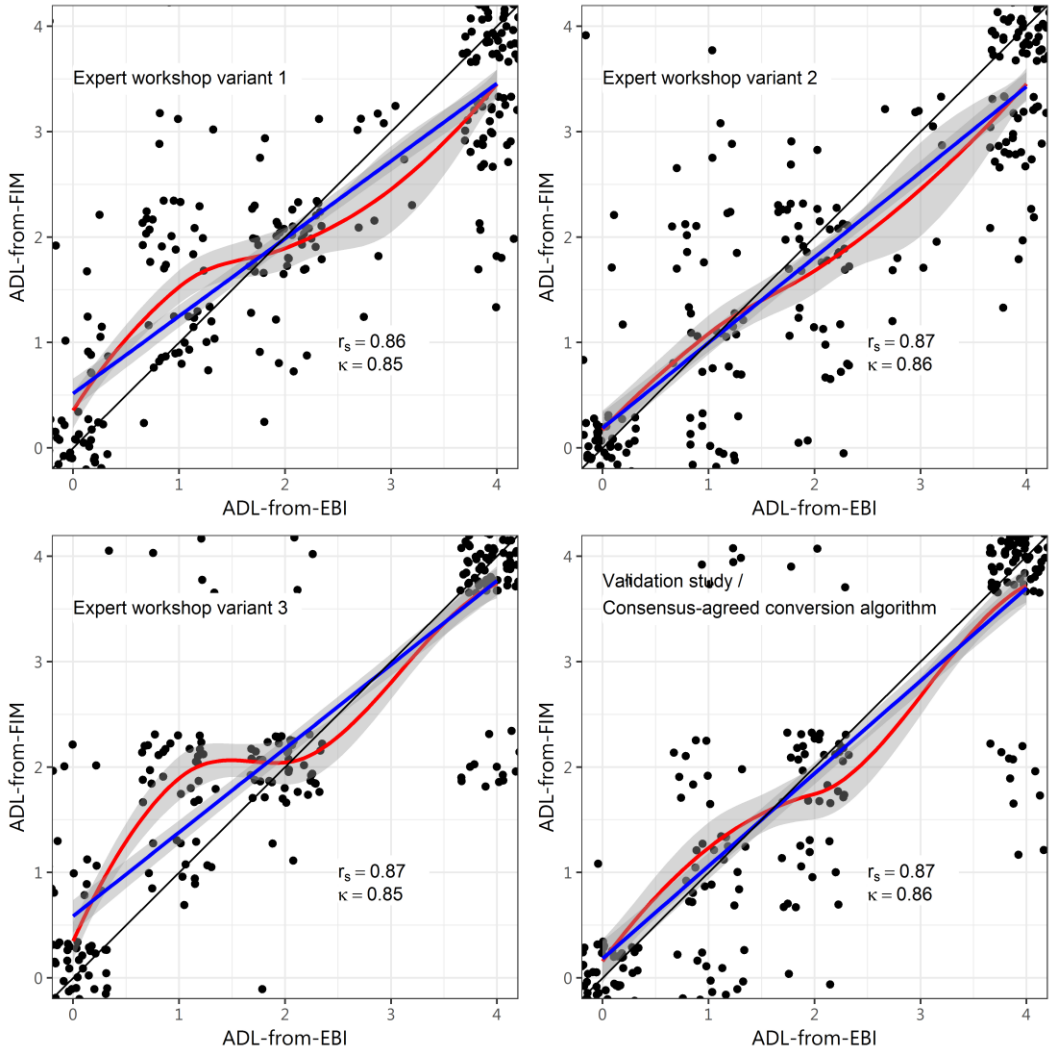


Table 61: Overview of conversion algorithm for item 2

Expert workshop variant 1		
EBI 2	**ADL**	**FIM® B**
0	0	1
1	1	2
		3
2	2	4
		5
3	3	6
4	4	7

Expert workshop variant 2		
EBI 2	**ADL**	**FIM® B**
0	0	1
		2
1	1	3
		4
2	2	5
3	3	6
4	4	7

Expert workshop variant 3		
EBI 2	**ADL**	**FIM® B**
0	0	1
1	1	2
2	2	3
		4
		5
3	4	6
4		7

Validation study		
EBI 2	**ADL**	**FIM® B**
0	0	1
		2
1	1	3
		4
2	2	5
3	4	6
4		7

Consensus-agreed conversion algorithm		
EBI 2	**ADL**	**FIM® B**
0	0	1
		2
1	1	3
		4
2	2	5
3	4	6
4		7

Item 3: Dressing and undressing

Figure 69: Scatter plot item 3 (Dressing and undressing)

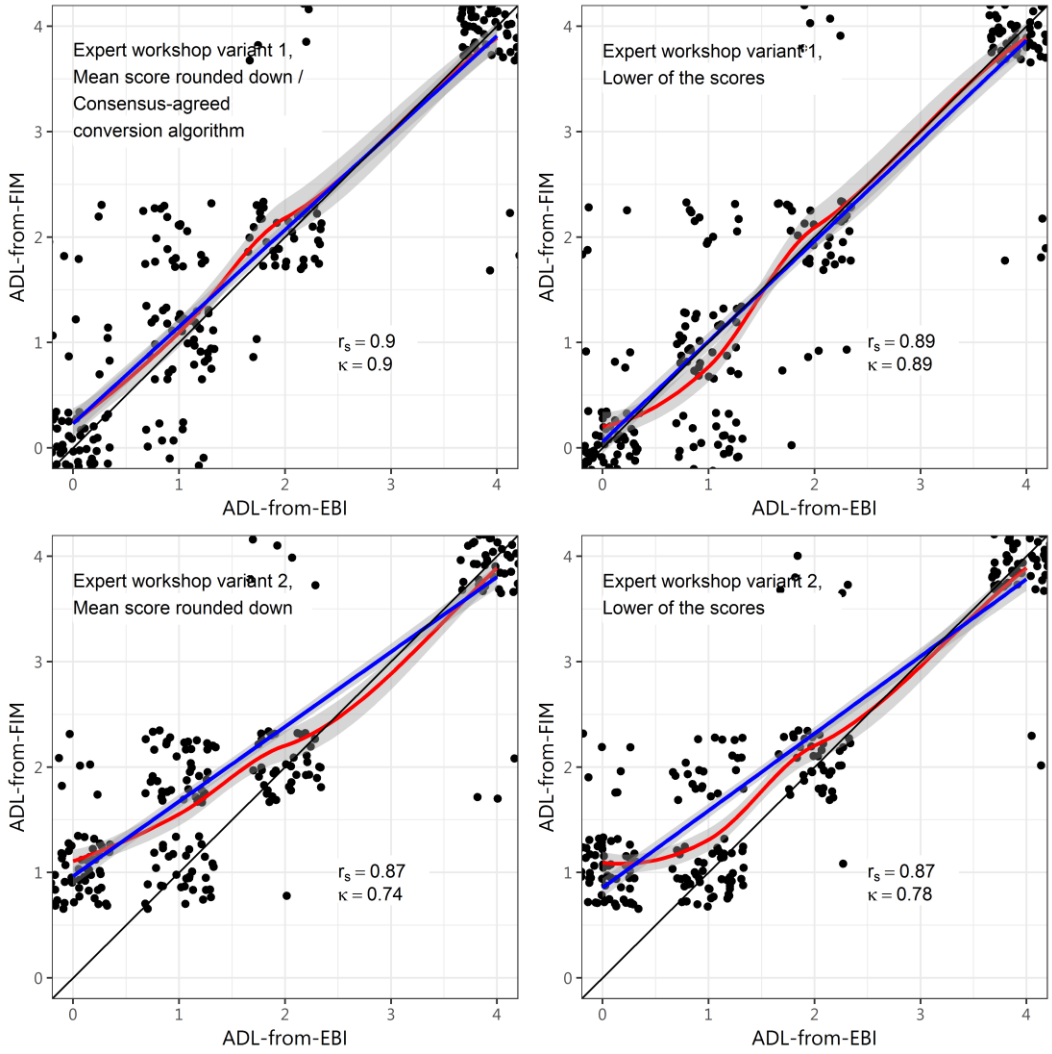


Table 62: Overview of conversion algorithm for item 3

Expert workshop variant 1		
EBI 3	**ADL**	**FIM® D+E***
0	0	1
1	1	2
		3
2	2	4
		5
4	4	6
		7

* Variant 1A: lower of the scores for D and E
* Variant 1B: mean score of D and E (rounded down, if not a whole number)

Expert workshop variant 2		
EBI 3	**ADL**	**FIM® D+E***
0	0	
1	1	1,2
2	2	3
		4
		5
4	4	6
		7

* Variant 2A: lower of the scores for D and E
* Variant 2B: mean score of D and E (rounded down if not a whole number)

Consensus-agreed conversion algorithm		
EBI 3	**ADL**	**FIM® D+E***
0	0	1
1	1	2
		3
2	2	4
		5
4	4	6
		7

* Mean score of D and E (rounded down if not a whole number)

Item 4: Bathing/showering/taking a complete sponge bath

Figure 70: Scatter plot item 4 (Bathing/showering/taking a complete sponge bath)

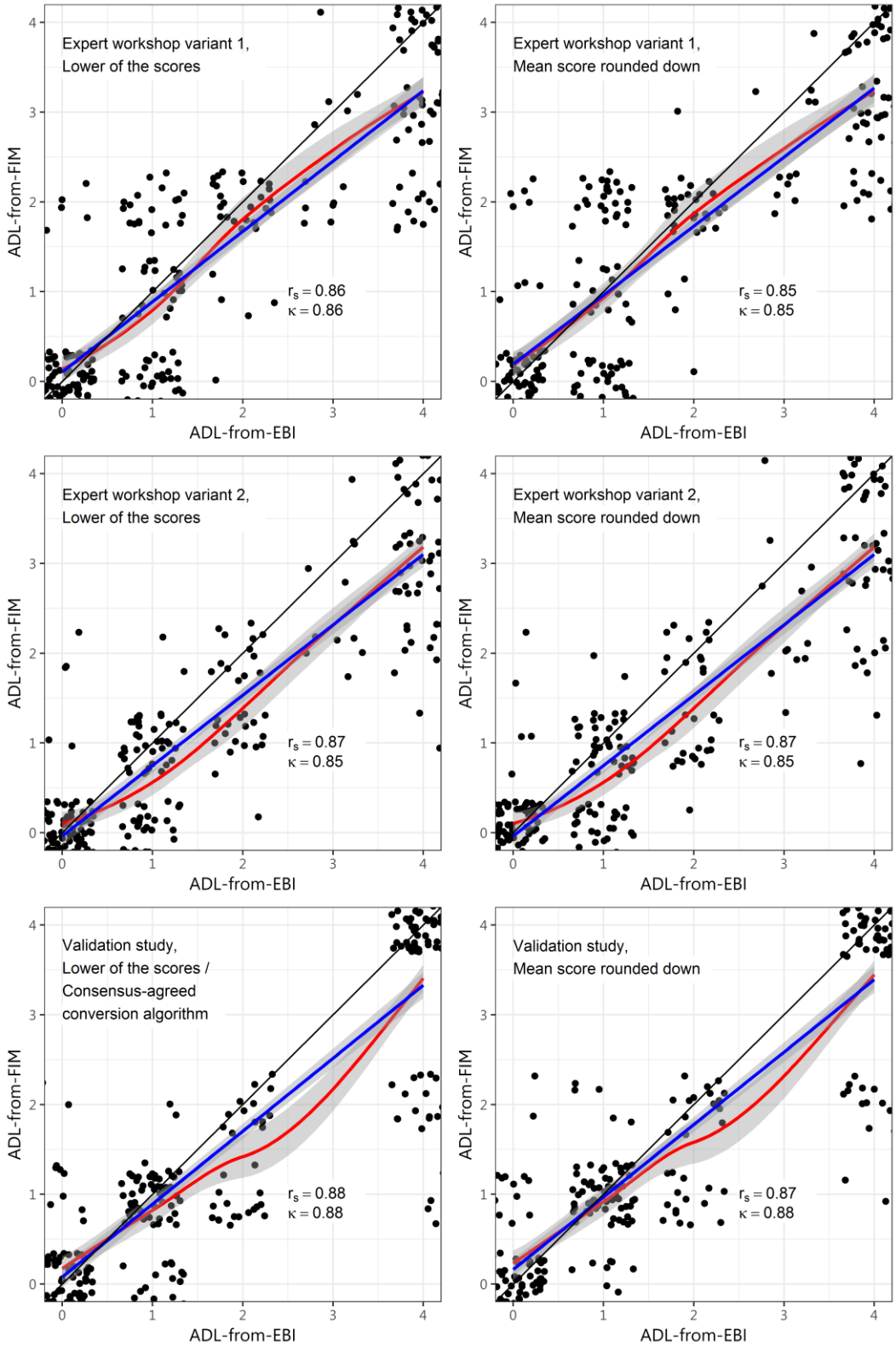


Table 63: Overview of conversion algorithm for item 4

Expert workshop variant 1		
EBI 4	**ADL**	**FIM® C+K***
0	0	1
		2
1	1	3
2	2	4
		5
3	3	6
4	4	7

* Variant 1A: lower of the scores for C and K
* Variant 1B: mean score of C and K (rounded down if not a whole number)

Expert workshop variant 2		
EBI 4	**ADL**	**FIM® C+K***
0	0	1
		2
1	1	3
		4
2	2	5
3	3	6
4	4	7

* Variant 2A: lower of the scores for C and K
* Variant 2B: mean score of C and K (rounded down if not a whole number)

Validation study		
EBI 4	**ADL**	**FIM® C+K***
0	0	1
1	1	2
		3
		4
2	2	5
3	4	6
4		7

* Variant VA: lower of the scores for D and E
* Variant VB: mean score of D and E (rounded down if not a whole number)

Consensus-agreed conversion algorithm		
EBI 4	**ADL**	**FIM® C+K***
0	0	1
1	1	2
		3
		4
2	2	5
3	4	6
4		7

* Variante VA: * mean score of C and K

Item 5: Bed to chair/wheelchair transfer

Figure 71: Scatter plot item 5 (Bed to chair/wheelchair transfer)

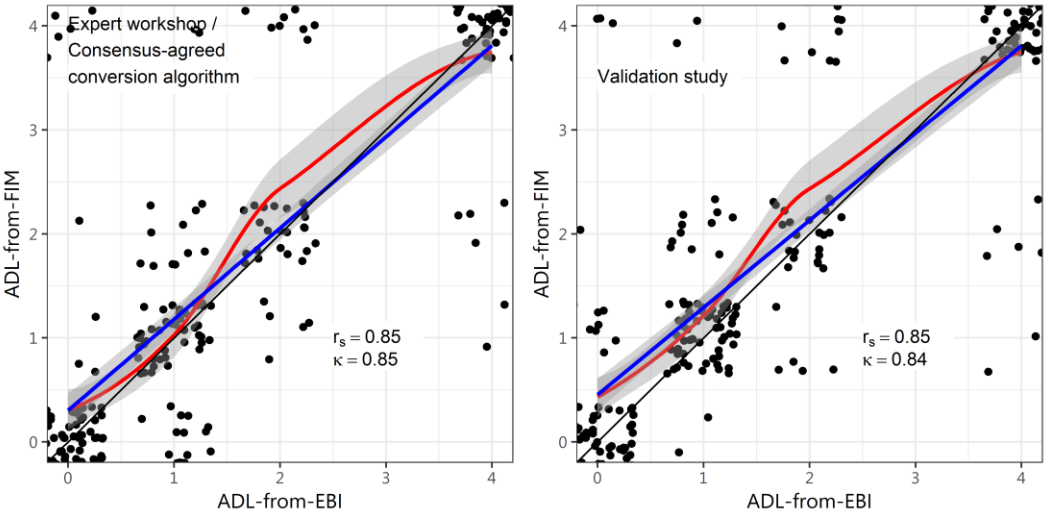


Table 64: Overview of conversion algorithm for item 5

Expert workshop / Consensus-agreed conversion algorithm		
EBI 5	**ADL**	**FIM® I**
0	0	1
		2
1	1	3
		4
2	2	5
4	4	6
		7

Validation study		
EBI 5	**ADL**	**FIM® I**
0	0	1
1	1	2
		3
		4
2	2	5
4	4	6
		7

Item 6: Locomotion

Figure 72: Scatter plot item 6 (Locomotion)

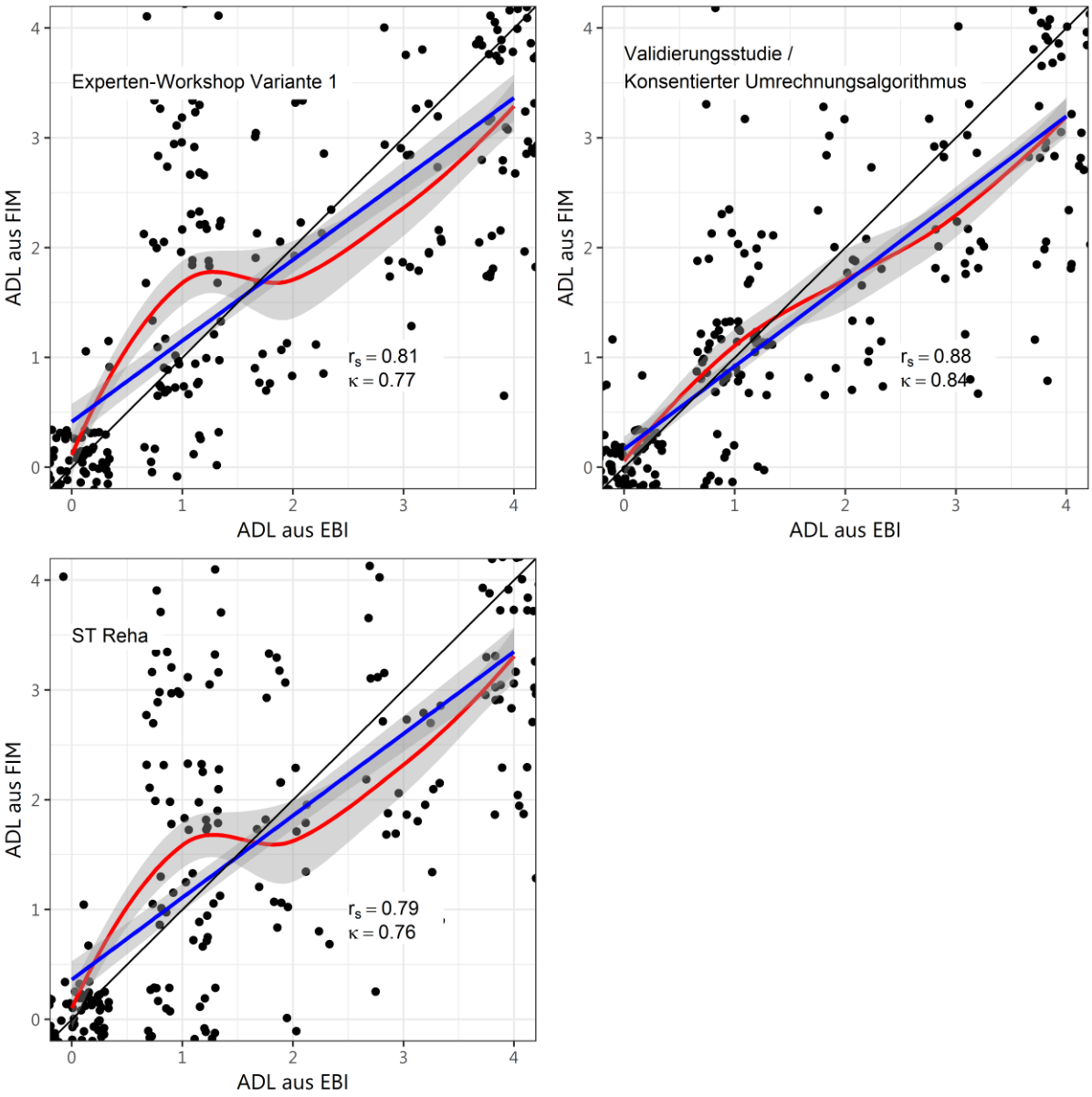


Table 65: Scheme for conversion algorithm for item 6

Expert workshop variant Walking Consensus-agreed conversion algorithms		
EBI 6	**ADL**	**FIM® L**
0	0	1
		2
1	1	3
		4
2	2	5
3	3	6
4	4	7

Expert workshop variant Wheelchair		
EBI 6	**ADL**	**FIM® L**
0	0	1
		2
1	1	3
		4
2	2	5
		6

Validation study Wheelchair Consensus-agreed conversion algorithms		
EBI 6	**ADL**	**FIM® L**
0	0	1
		2
1	1	3
		4
2		5
3		6
4		7

Validation study Walking AND Wheelchair Consensus-agreed conversion algorithms		
EBI 6	**ADL**	**FIM® L**
0	0	1
		2
1	1	3
		4
2	2	5
3		6
4		7

Item 7: Stairs

Figure 73: Scatter plot item 7 (Stairs)

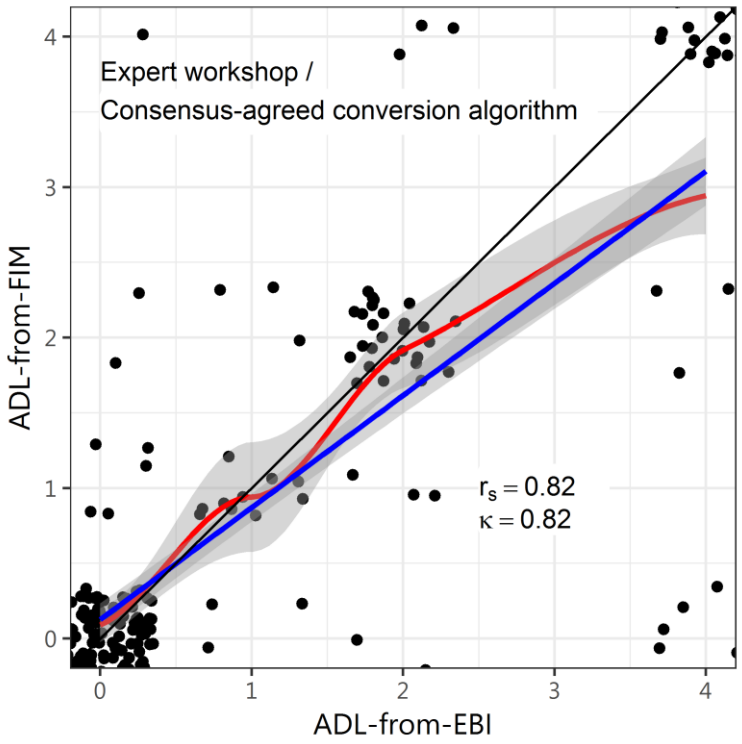


Table 66: Overview of conversion algorithm for item 7

Expert workshop		
EBI 7	**ADL**	**FIM® M**
0	0	1
1	1	2 3
2	2	4 5
4	4	6 7

Consensus-agreed conversion algorithms		
EBI 7	**ADL**	**FIM® M**
0	0	1
1	1	2 3
2	2	4 5
4	4	6 7

Item 8: Toilet use

Figure 74: Scatter plot item 8 (Toilet use)

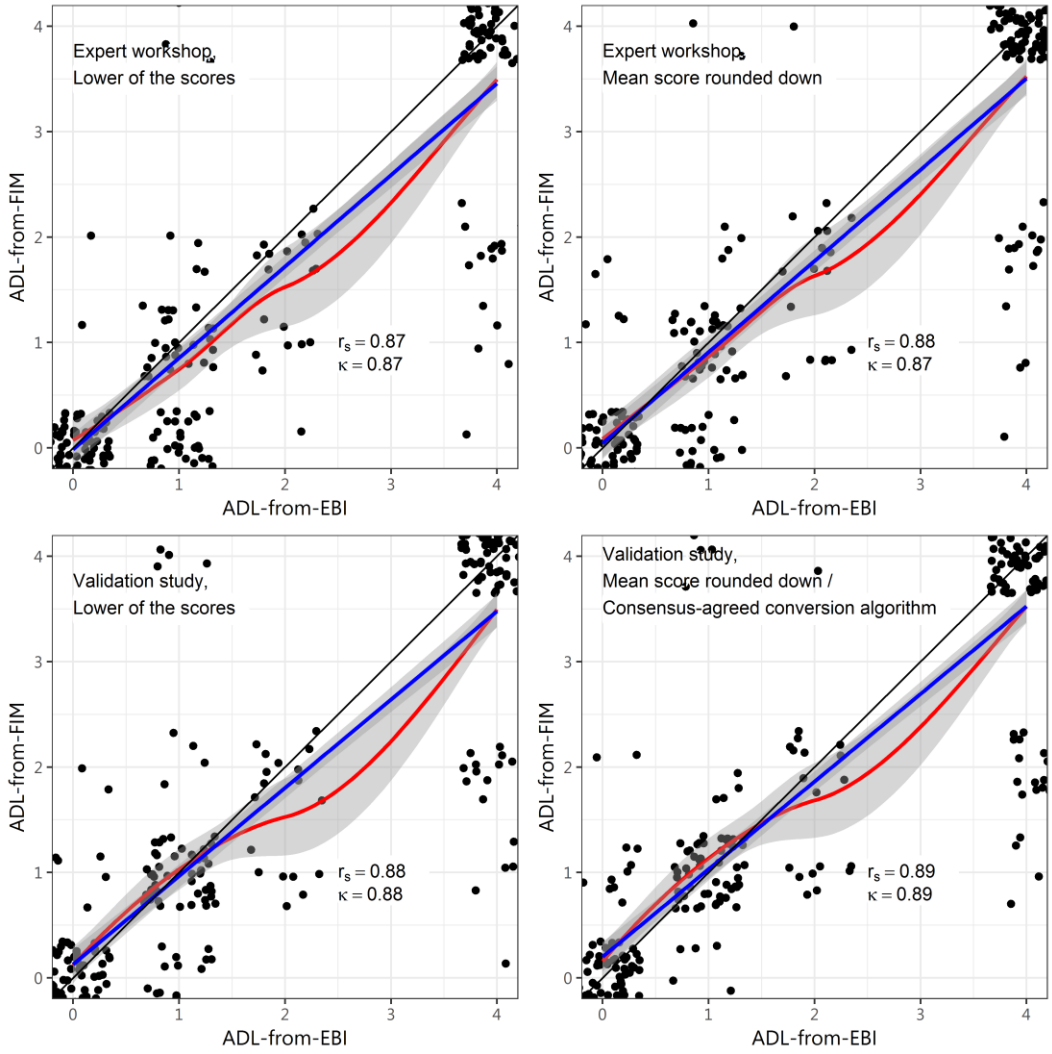


Table 67: Overview of conversion algorithm for item 8

Expert workshop		
EBI 8	**ADL**	**FIM® F+J***
0	0	1
		2
1	1	3
		4
2	2	5
4	4 (0**)	6
		7

* Variant 1A: lower of the two values for F and J
* Variant 1B: Mean score of F and J (rounded down if not a whole number)
** If EBI 5 (Bed to chair/wheelchair transfer) and EBI item 6 (Locomotion) are each scored as "0", the ADL score "4" for item 8 is revised to "0".

Validation study		
EBI 8	**ADL**	**FIM® F+J***
0	0	1
1	1	2
		3
		4
2	2	5
4	4 (0**)	6
		7

* Variant 1A: lower of the scores of F and J
* Variant 1B: mean score of F and J (rounded down if not a whole number)
** if the scores of EBI 5 (Transfer: wheelchair / bed & vv.) and EBI Item 6 (Locomotion on level surfaces) are "0" each, the ADL-Score "4" of Item 8 is set to "0" as well.

Consensus-agreed conversion algorithms		
EBI 8	**ADL**	**FIM® F+J***
0	0	1
1	1	2
		3
		4
2	2	5
4	4 (0**)	6
		7

* Variant 1B: mean score of F and J (rounded down if not a whole number)
** if the scores of EBI 5 (Transfer: wheelchair / bed & vv.) and EBI Item 6 (Locomotion on level surfaces) are "0" each, the ADL-Score "4" of Item 8 is set to "0" as well.

Item 9: Bowel management

Figure 75: Scatter plot item 9 (Bowel management)

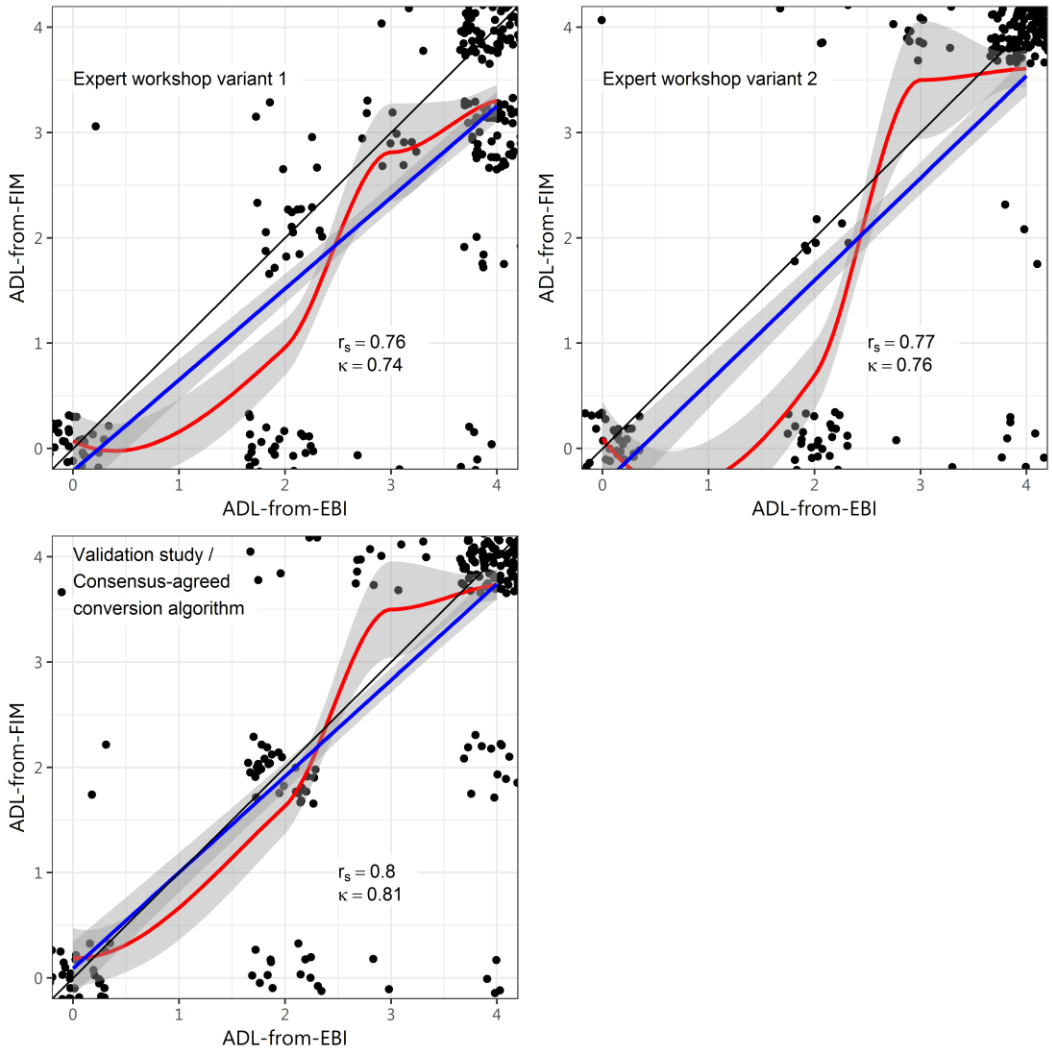


Table 68: Overview of conversion algorithm for item 9

Expert workshop - variant 1		
EBI 9	ADL	FIM® H
0	0	1
		2
2	2	3
		4
3	3	5
		6
4	4	7

Expert workshop - variant 2		
EBI 9	ADL	FIM® H
0	0	1
		2
		3
2	2	4
3	4	5
		6
4		7

Validation study		
EBI 9	ADL	FIM® H
0	0	1
2	2	2
		3
		4
3	4	5
4		6
		7

Consensus-agreed conversion algorithms		
EBI 9	ADL	FIM® H
0	0	1
2	2	2
		3
		4
3	4	5
4		6
		7

Item 10: Bladder management

Figure 76: Scatter plot item 10 (Bladder management)

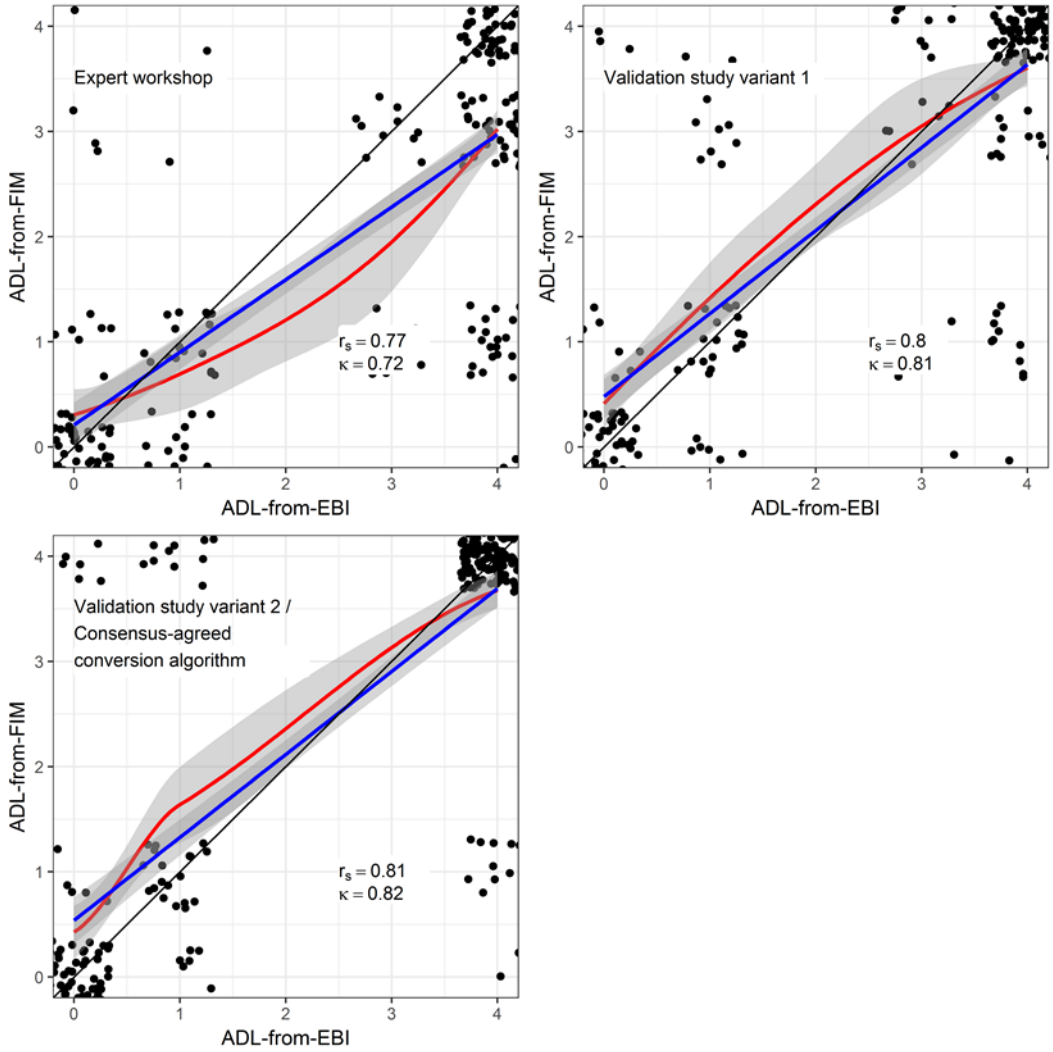


Table 69: Overview of conversion algorithm for item 10

Expert workshop		
EBI 10	**ADL**	**FIM® G**
0	0	1
		2
1	1	3
		4
		5
3	3	6
4	4	7

Validation study variant 1		
EBI 10	**ADL**	**FIM® G**
0	0	1
1	1	2
		3
		4
	3	5
3	4	6
4		7

Validation study variant 2		
EBI 10	**ADL**	**FIM® G**
0	0	1
1	1	2
		3
		4
3	4	5
4		6
		7

Consensus-agreed conversion algorithms		
EBI 10	**ADL**	**FIM® G**
0	0	1
1	1	2
		3
		4
3	4	5
4		6
		7

Item 11: Comprehension

Figure 77: Scatter plot item 11 (Comprehension)

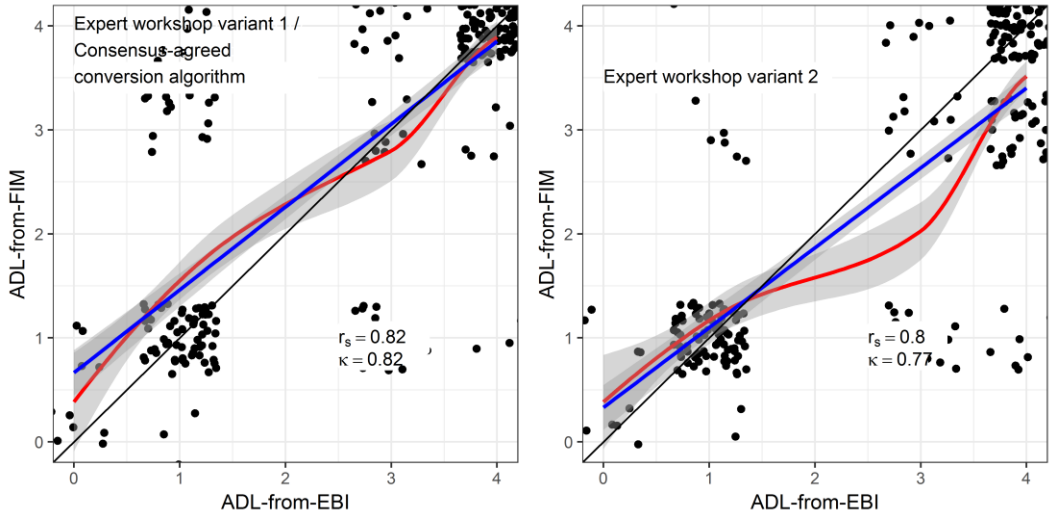


Table 70: Overview of conversion algorithm for item 11

Expert workshop variant 1		
EBI 11	**ADL**	**FIM® N**
0	0	1
1	1	2
		3
		4
3	3	5
4	4	6
		7

Expert workshop variant 2		
EBI 11	**ADL**	**FIM® N**
0	0	1
1	1	2
		3
		4
		5
3	3	6
4	4	7

Consensus-agreed conversion algorithms		
EBI 11	**ADL**	**FIM® N**
0	0	1
1	1	2
		3
		4
3	3	5
4	4	6
		7

Item 12: Expression

Figure 78: Scatter plot item 12 (Expression)

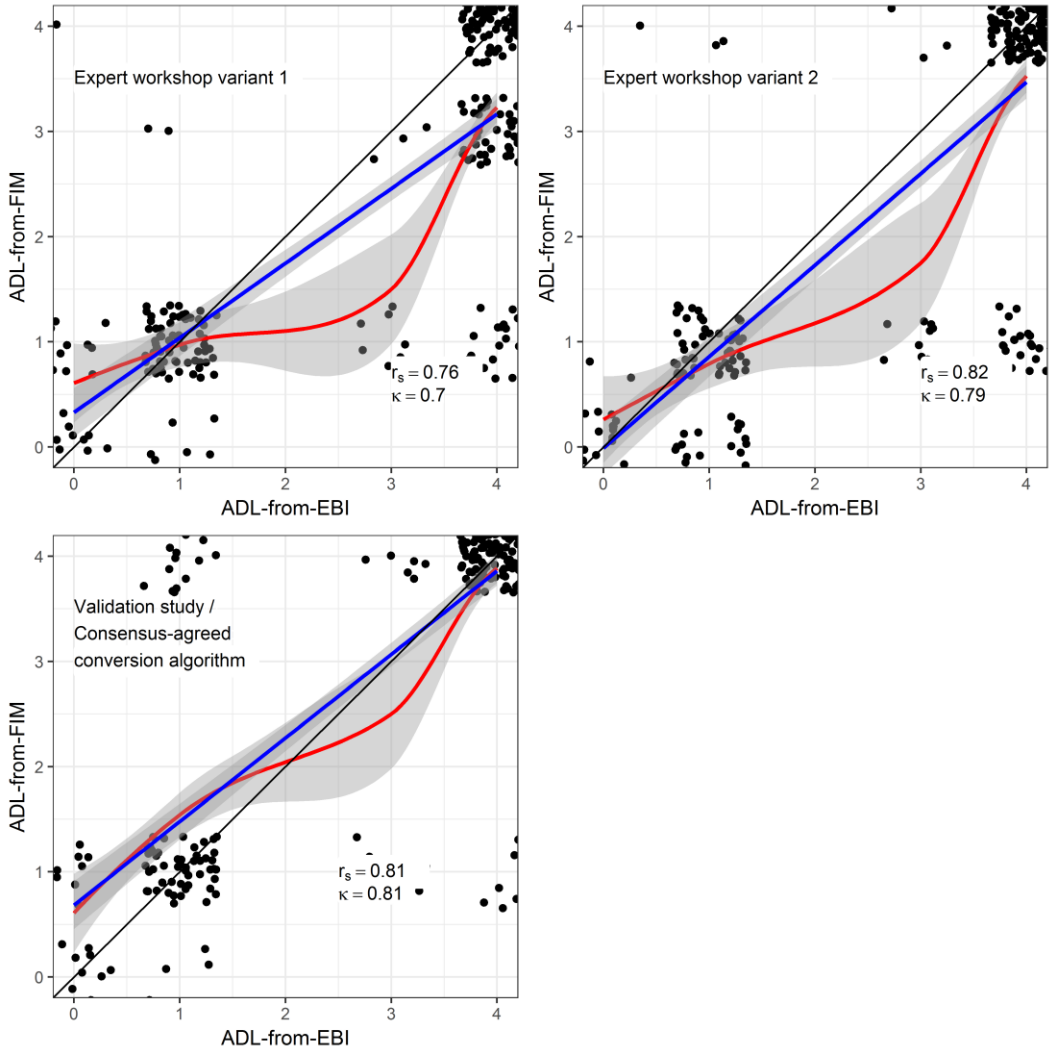


Table 71: Overview of conversion algorithm for item 12

Expert workshop variant 1		
EBI 12	**ADL**	**FIM® O**
0	0	1
1	1	2
		3
		4
		5
3	3	6
4	4	7

Expert workshop variant 2		
EBI 12	**ADL**	**FIM® O**
0	0	1
		2
1	1	3
		4
		5
3	4	6
4		7

Validation study		
EBI 12	**ADL**	**FIM® O**
0	0	1
1	1	2
		3
		4
3	4	5
4		6
		7

Consensus-agreed conversion algorithms		
EBI 12	**ADL**	**FIM® O**
0	0	1
1	1	2
		3
		4
3	4	5
4		6
		7

Item 13: Social interaction

Figure 79: Scatter plot item 13 (Social interaction)

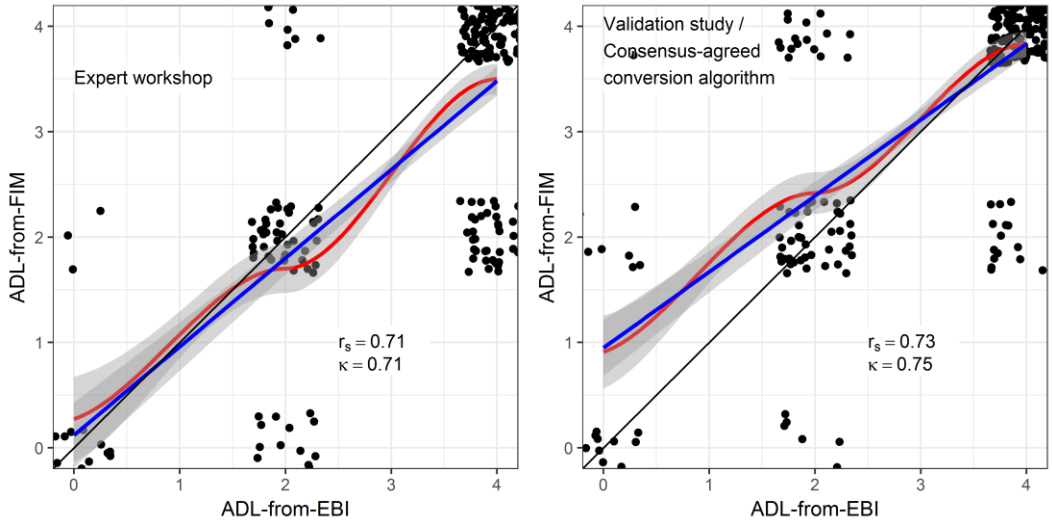


Table 72: Overview of conversion algorithm for item 13

Expert workshop		
EBI 13	**ADL**	**FIM® P**
0	0	1
		2
2	2	3
		4
		5
4	4	6
		7

Validation study		
EBI 13	**ADL**	**FIM® P**
0	0	1
		2
2	2	3
		4
		5
4	4	6
		7

Consensus-agreed conversion algorithms		
EBI 13	**ADL**	**FIM® P**
0	0	1
		2
2	2	3
		4
		5
4	4	6
		7

Item 14: Problem solving

Figure 80: Scatter plot item 14 (Problem solving)

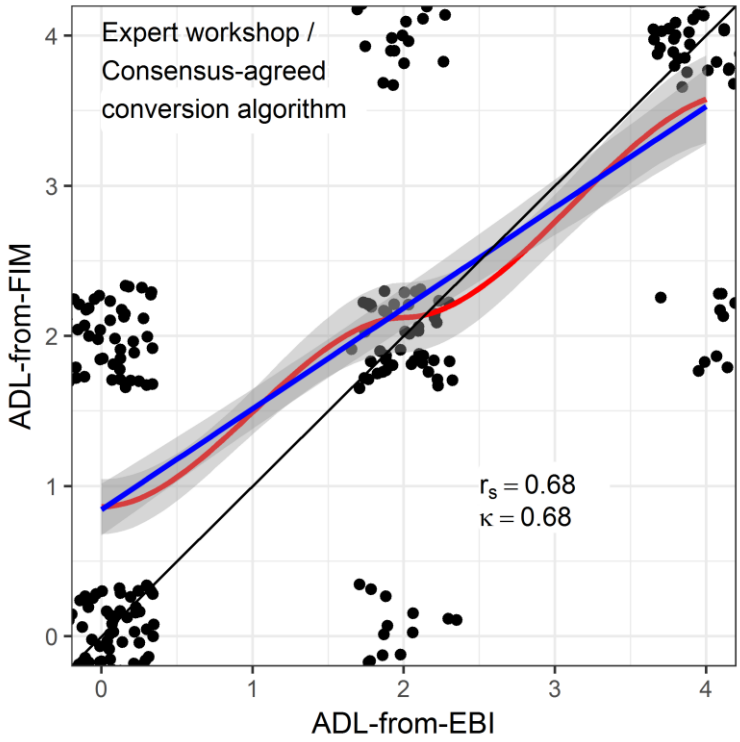


Table 73: Overview of conversion algorithm for item 14

Expert workshop		
EBI 14	**ADL**	**FIM® Q**
0	0	1 2
2	2	3 4 5
4	4	6 7

Consensus-agreed conversion algorithms		
EBI 14	**ADL**	**FIM® Q**
0	0	1 2
2	2	3 4 5
4	4	6 7

Item 15: Memory/learning capability/orientation

Figure 81: Scatter plot item 15 (Memory/learning capability/orientation)

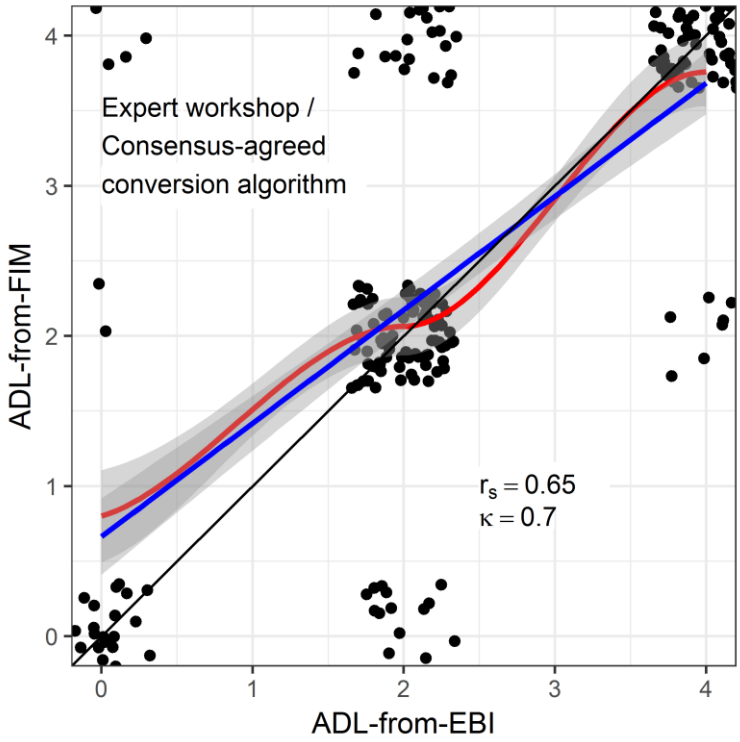


Table 74: Overview of conversion algorithm for item 15

Expert workshop		
EBI 15	**ADL**	**FIM® R**
0	0	1
1		2
2	2	3
3		4
		5
4	4	6
		7

Consensus-agreed conversion algorithms		
EBI 15	**ADL**	**FIM® R**
0	0	1
1		2
2	2	3
3		4
		5
4	4	6
		7

A3 Consensus-approved conversion algorithms for all items

Table 75: Consensus-agreed algorithms for the conversion of EBI and FIM scores to ADL scores for all items

EBI Item	ADL Score	FIM® Item	explanatory remark
1. Eating	**ADL 1**	**A. Eating**	
0	0	1 - 2	
2	2	3 - 5	
3 - 4	4	6 - 7	
2. Grooming	**ADL 2**	**B. Grooming**	
0	0	1 - 2	
1	1	3 - 4	
2	2	5	
3 -4	4	6 - 7	
3. Dressing / undressing	**ADL 3**	**D. Dressing - Upper Body / E. Dressing - Lower Body***	* The mean score of Items FIM® D. and Items FIM® E. is used and rounded down if it's not a whole number.
0	0	1	
1	1	2 - 3	
2	2	4 - 5	
4	4	6 - 7	
4. Bathing / showering / taking a complete sponge bath	**ADL 4**	**C. Bathing K. Transfer: Tub / Shower***	* The lower of the scores of Items FIM® C. and FIM® K. is used.
0	0	1	
1	1	2 - 4	
2	2	5	
3 - 4	4	6 - 7	
5. Transfer: wheelchair / bed & vv.	**ADL 5**	**I. Transfer: Bed / Chair / Wheelchair**	
0	0	1 - 2	
1	1	3 - 4	
2	2	5	
4	4	6 - 7	
6. Locomotion on level surfaces	**ADL 6**	**L. Locomotion: Walk / Wheelchair***	
0	0	1 - 2	If additional Item Walking was used.
1	1	3 - 4	
2	2	5	
3	3	6	
4	4	7	
0	0	1 - 2	If additional Item Wheelchair was used.
1 - 4	1	3 - 7	
0	0	1 - 2	If additional Item Walking AND Wheelchair was used.
1	1	3 - 4	
2 - 4	2	5 - 7	
7. Ascending / descending stairs	**ADL 7**	**M. Locomotion: Stairs**	
0	0	1	
1	1	2 - 3	
2	2	4 - 5	
4	4	6 - 7	

... continued on next page.

EBI Item	ADL Score	FIM® Item	Bemerkung
8. Use of the toilet	**ADL 9**	**F. Toileting** **J. Transfer: Toilet***	* * The mean score of Items FIM® F. and Items FIM® J. is used and rounded down if it's not a whole number. ** if the scores of EBI 5 (Transfer: wheelchair / bed & vv.) and EBI Item 6 (Locomotion on level surfaces) are "0" each, the ADL-Score "4" of Item 8 is set to
0	0	1	
1	1	2 - 4	
2	2	5	
4	4 (0**)	6 - 7	
9. Bowel management	**ADL 9**	**H. Bowel Management**	
0	0	1	
2	2	2 - 4	
3 - 4	4	5 - 7	
10. Bladder Management	**ADL 10**	**G. Bladder Management**	
0	0	1	
1	1	2 - 4	
3 - 4	4	5 - 7	
11. Cognitive comprehension	**ADL 11**	**N. Comprehension**	
0	0	1	
1	1	2 - 4	
3	3	5	
4	4	6 - 7	
12. Expression	**ADL 12**	**O. Expression**	
0	0	1	
1	1	2 - 4	
3 - 4	4	5 - 7	
13. Social Interaction	**ADL 13**	**P. Social Interaction**	
0	0	1	
2	2	2 - 4	
4	4	5 - 7	
14. Problem Solving	**ADL 14**	**Q. Problem Solving**	
0	0	1 - 2	
2	2	3 - 5	
4	4	6 - 7	
15. Memory / learning / orientation	**ADL 15**	**Q. Memory**	
0 - 1	0	1 - 2	
2 - 3	2	3 - 5	
4	4	6 - 7	

Imprint

Title	Development and validation of an algorithm to convert FIM® and EBI to an ADL score
Authors	Martin Brünger, MPH Dipl.-Päd. (Rehab.) Stefanie Köhn Dr. Anna Schlumbohm Prof. Karla Spyra
Place and year of publication	Berlin, 2 March 2018
Sponsor	Swiss National Association for Quality Development in Clinics and Hospitals – ANQ
represented by	Dr. Luise Menzi, Head of Rehabilitation

Registered office

Weltpoststrasse 5, Postfach 370

CH-3015 Bern

Charité – Universitätsmedizin Berlin

Institute of Medical Sociology and Rehabilitation Science

Charitéplatz 1

D-10117 Berlin