

1 **Rasch validation of a new scale to measure dependency in arm use in**
2 **daily life: the Upper Limb Lucerne ICF-based Multidisciplinary**
3 **Observation Scale (UL-LIMOS)**

4 **Ann Van de Winckel, PhD, MSPT, PT^{1*}, Beatrice Ottiger, MSc², Janne Marieke Veerbeek,**
5 **PhD², Thomas Nyffeler, MD^{2,3,4}, Tim Vanbellingen, PhD^{2,3}**

6 ¹Brain Body Mind Laboratory, Division of Physical Therapy, Division of Rehabilitation Science,
7 Department of Rehabilitation Medicine, Medical School, University of Minnesota

8 ²Neurocenter, Luzerner Kantonsspital, Luzern

9 ³ARTORG Center for Biomedical Engineering Research, Gerontechnology and Rehabilitation
10 Group, University of Bern

11 ⁴Departement of Neurology, Inselspital, Bern University Hospital, University of Bern, Bern,
12 Switzerland

13 *** Correspondence:**

14 Corresponding Author: Ann Van de Winckel, PhD, MSPT, PT
15 avandewi@umn.edu

16 **Keywords:** Outcome; Rasch Measurement Theory; Rasch Analysis; Upper limb; Disability; Health;
17 Stroke; Activities of Daily Living

18 **Word count: 3021**

19 **Number of Figures: 1**

20 **Number of Tables: 4**

21 **Abstract**

22 **Introduction:** About 77% of adults with stroke have upper limb impairments. Many upper limb
23 measures are available for adults with stroke to measure the impairment and activity level of the
24 affected limb. However, an observational scale focused on assessing dependency in upper limb use
25 during daily life activities (as opposed to testing in laboratory settings) is lacking. To bridge this gap,
26 we have developed a new 5-item “Upper Limb Lucerne ICF-based Multidisciplinary Observation
27 Scale (UL-LIMOS)”, which assesses dependency on others during affected arm use in daily life in
28 adults with stroke. As a next step in the psychometric analysis, we evaluated the unidimensionality
29 and structural validity of the UL-LIMOS with Rasch Measurement Theory.

30 **Methods:** This is a single-center cross-sectional study in adults with (sub)acute stroke. We applied
31 Rasch Measurement Theory (RMT) to analyze the structural validation and unidimensionality of the
32 new UL-LIMOS. We chose a polytomous partial credit model using the Rasch Unidimensional
33 Measurement Model (RUMM) 2030 software. The outputs provide evidence of unidimensionality,
34 item and person fit, overall fit, principal component analysis of residuals (PCAR), person separation
35 reliability (PSR), as well as residual item correlations to identify local item dependence. Person mean
36 location, floor and ceiling effects identify proper targeting.

37 **Results:** We recruited 407 adults with (sub)acute stroke (median age 63 years, 157 women). All
38 items and persons fit the Rasch model, and the PSR of 0.90 indicates that clinicians and researchers
39 can reliably use the scale for individual decision-making. There were small floor (2.70%) and ceiling
40 (13.00%) effects. The average person mean location was 1.32 ± 2.99 logits, indicating that the items
41 were too easy for this group of adults with (sub)acute stroke. The PCAR’s eigenvalue was 2.46 with
42 49.23% explained variance. Further analysis of paired *t*-tests revealed that 0.89% of person locations
43 were significantly different when comparing the two subtests formed based on positive and negative
44 loadings on the first principal component, thereby confirming the unidimensionality of the scale. One
45 pair of items related to “arm and hand use” and “fine hand use” showed residual item correlations.

46 **Discussion:** The new Rasch-based UL-LIMOS is a valid ICF-based observation scale at the ICF-
47 participation level, to evaluate dependency during upper limb use in daily life in adults with stroke.
48 The UL-LIMOS would be a valuable addition to the core assessments of adults with (sub)acute
49 stroke in hospitals and rehabilitation centers. Further analysis is needed to generalize our findings to
50 adults with chronic stroke who have returned to their home setting, and in other countries to account
51 for cultural differences. Targeting could be improved in the future. Additional psychometric analyses,
52 such as sensitivity to change, are warranted. A comparison of the UL-LIMOS data with self-reported
53 measurements or accelerometers could potentially lead to changes to the core datasets recommended
54 for the evaluation of adults with stroke.

55

56 1 Introduction

57 About 77% of adults with stroke have upper limb impairments (1). These impairments hinder
58 performing activities of daily life (ADL) independently and result in long-term dependency in 50%
59 of the cases (2,3). This dependency decreases the quality of life (QoL) (4) and results in an inability
60 to return to work in 40% of working-age adults who have had a stroke (5).

61 Many upper limb measures evaluate motor recovery after stroke (for an overview see (6)). Following
62 the International Classification of Functioning, Disability and Health (ICF) (7), these measures assess
63 either upper limb impairment (e.g., mobility of joints, muscle power/tone/endurance), or assess
64 activities (e.g., grasping a block of wood) in laboratory settings.

65 Of all the upper limb measures, the Upper Extremity Subscale of the Fugl-Meyer Motor Assessment
66 (FMA-UE) (8) on the impairment level and the Action Research Arm Test (ARAT) (9) on the
67 activity level were suggested as core upper limb assessments for stroke rehabilitation trials (10) and
68 clinical rehabilitation (11). So far, there is no consensus on scale recommendations for assessing
69 dependency in upper limb use in daily life (10). The Barthel Index (BI) (12) and the Functional
70 Independence Measure (FIM) (13) –originally designed to evaluate the need for nursing care– are
71 commonly used in clinical rehabilitation and research to assess daily life functioning in general
72 (11,14), such as feeding, bathing, dressing, and undressing. However, BI (15) and FIM (16) do not
73 focus on specific upper limb use in daily life. Moreover, problematic floor and ceiling effects (>15%)
74 have been reported (15–17). Some patient-reported outcome measures, such as the Motor Activity
75 Log (MAL) (18), or ABILHAND (19), which are administered through semi-structured interviews,
76 were developed to evaluate the stroke individual’s perspective on real-life upper limb performance.
77 Yet, due to the subjective nature of patients’ reports, these measures should not be used with adults
78 with stroke who have moderate to severe cognitive deficits. For this reason, they were not suggested
79 as core measures to assess upper limb performance in daily life (10). The Actual Amount of Use Test
80 (AAUT) is an observational test, which measures how much patients spontaneously use their affected
81 arm during predefined tasks in a laboratory setting. Therefore, the AAUT also does not reflect
82 spontaneous upper limb use in daily life (20).

83 Others have used accelerometers as a measure of upper limb performance in daily life, which has the
84 advantage of not being biased by patients’ subjective reports (21,22). However, accelerometers
85 cannot determine what type of activity was performed. In sum, an observational scale specifically
86 focused on assessing dependency in upper limb use during actual daily life activities (as opposed to
87 testing in a laboratory setting) is lacking.

88 Recently, we demonstrated that the Lucerne ICF-based Multidisciplinary Observation Scale
89 (LIMOS), a clinician-reported measure, was reliable and valid in evaluating the performance of
90 activities in daily life in adults with acute and subacute stroke (23–26). Clinician-reported measures
91 are measures scored by a health care professional based on observing the patients’ spontaneous
92 behaviors, e.g., during their stay in the hospital or rehabilitation center. It can therefore be used in
93 adults with moderate to severe cognitive impairments after stroke. LIMOS covered several domains
94 (motor, communication, learning and applying knowledge, and domestic life), and showed no
95 problematic floor or ceiling effects (23,26). Moreover, responsiveness was higher for LIMOS than
96 for BI and FIM (25).

97 Based on this previous work, and to fill the gap concerning assessing actual arm and hand use in
98 daily life after a stroke, we developed a new, 5-item upper Limb Lucerne ICF-based
99 Multidisciplinary Observation Scale (UL-LIMOS), which evaluates upper limb use in daily life. The
100 goal of developing this new measure, derived from the reliable and valid LIMOS (23–26), is to
101 obtain a quick evaluation measurement of dependency on others for upper limb use in daily life, for

102 use in the hospital and rehabilitation centers, or in research (27). As a next step in establishing the
103 psychometrics of this scale, we aim to test the structural validity and unidimensionality of the UL-
104 LIMOS, using Rasch measurement Theory.

105

106 **2 Materials and methods**

107 **2.1 Participants**

108 We approached adults with (sub)acute stroke for participation in the study, who were admitted for
109 inpatient neurorehabilitation in the rehabilitation center Neurocenter, Luzerner Kantonsspital,
110 Lucerne, in Switzerland. Stroke diagnosis was based on the European Stroke Organization (ESO)
111 guidelines, which are based on both clinical and Magnetic Resonance Imaging (MRI) criteria. We
112 included adults with a first-ever acute to subacute stroke, up to 6 months post-stroke, showing
113 unilateral ischemic or hemorrhagic supratentorial lesions (28). Adults with stroke were excluded if
114 they had bilateral lesions. There were no other inclusion and exclusion criteria.

115 The study was conducted in accordance with the principles of the Declaration of Helsinki (2013) and
116 was approved by the local Ethical Committees of the state of Luzern (BASEC-ID 2017-00998). We
117 followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)
118 Statement (29). The participants gave written informed consent. A family member with power of
119 attorney consented if participants had severe cognitive impairments, preventing them to consent
120 independently.

121 **2.2 Main outcome measures**

122 We acquired demographic (sex, age) and clinical data (type and time after stroke, location of stroke,
123 presence of cognitive deficits (30–32), and more specifically, unilateral neglect (33) and apraxia (34–
124 36) for their impact on motor function) as well as the UL-LIMOS at admission. The UL-LIMOS is
125 composed of 5 items, which are items selected from the more encompassing LIMOS (23–26). The
126 LIMOS addressed the dependency of others during daily activities on several domains, among which
127 motor activities (with 18 items).

128 Our previous reliability and validity studies on LIMOS demonstrated high internal consistency
129 (coefficient $\alpha=0.98$), good test-retest reliability at the item level (moderate to excellent range of
130 kappa between 0.41 and 0.84, except for two items with fair agreement, kappa = 0.32–0.37), and
131 subscale levels (intraclass correlation coefficient $r>0.75$, range 0.76–0.95) (23,26). Inter-rater
132 reliability demonstrated moderate to excellent agreement with kappa values ranging from 0.41 to
133 0.92 (23,26) except for 12 items demonstrating fair agreement.

134 We demonstrated a strong convergent validity between LIMOS motor and FIM motor ($r = 0.89$),
135 between LIMOS motor and Barthel Index ($r = 0.92$), between LIMOS motor and FIM mobility ($r =$
136 0.90), and between the subscales LIMOS knowledge and FIM cognition ($r = 0.81$) (23,26).
137 Correlations between other subscales of the LIMOS (self-care, general tasks, domestic life) and the
138 subscales of the FIM ranged between $r = 0.36$ –0.79 (23,26). A moderate positive correlation was
139 found between LIMOS cognition and communication subscale and FIM cognition ($r = 0.67$)
140 (23,25,26).

141 The LIMOS motor subscale, and the applying knowledge and communication subscale were more
142 responsive, expressed by higher effect sizes (ES = 0.65, Standardized Response Mean, SRM = 1.17

143 and ES = 0.52, SRM = 1.17, respectively) as compared with FIM motor (ES = 0.54, SRM = 0.96)
144 and FIM cognition (ES = 0.41, SRM = 0.88) or Barthel (ES = 0.41, SRM = 0.65) (25).

145 Rasch-based LIMOS subscales fit the Rasch model after reducing and rescaling items: LIMOS
146 subscales motor (18 items), communication (5 items), applying knowledge/cognition (13 items), and
147 domestic life (5 items) (24).

148 Regarding **UL-LIMOS**, the 5 items, used in previous studies(27,37)are “*lifting and carrying objects*”
149 (item 1), “*fine hand use*” (item 2), “*hand and arm use*” (item 3), “*washing oneself*” (item 4), and
150 “*dressing*” (item 5). The items were scored on a 5-point scale (0 to 4) with 0 being “patient is not
151 able to fulfill a task or needs assistance up to 75% (corresponding to “complete”)”; 1 representing
152 “patient is able to fulfill tasks with assistance of 25% to 75% (corresponding to “severe”); 2 being
153 “patient is able to fulfill tasks with assistance less than 25% or under supervision (corresponding to
154 “moderate”)”; 3 representing “patient is able to fulfill tasks independently but needs more time
155 and/or with auxiliary materials, aids (corresponding to “slight”)”; and 4 being “patient is able to
156 fulfill tasks independently (corresponding to “none”)”. The 5 items are summed to obtain a total UL-
157 LIMOS score, ranging from 0 representing no use of the upper arm in daily life to 20 representing
158 independent use of the upper limb in daily life (for the manual containing the scoring sheet and
159 instructions, see *Supplementary Material*).

160 We previously demonstrated a strong positive correlation ($r=0.78$) between UL-LIMOS and handgrip
161 strength (assessed with the Jamar dynamometer), and a moderate negative correlation ($r=0.69$) with
162 the Catherine Bergego Scale, which quantifies the influence of spatial neglect-related deficits on the
163 ADL (37).

164 **2.3 Statistical analysis**

165 The Rasch Measurement Theory (RMT) was applied to analyze the structural validation and
166 unidimensionality of the new UL-LIMOS. We chose a polytomous partial credit model using the
167 Rasch Unidimensional Measurement Model (RUMM) 2030 software.

168 A major advantage of RMT is that UL-LIMOS items can be hierarchically ordered from easy to
169 difficult and that ordinal scales are converted to interval scales where person and item distributions
170 are located on the same ruler with logit units. The Rasch prediction model states that a person with a
171 higher level of independence in upper limb use in daily life has a higher probability of scoring higher
172 on items than a person with less independence in upper limb use in daily life.

173 We followed the RULER guideline recommendations for reporting Rasch-based studies, recently
174 published by Van de Winckel et al. (2022); Mallinson et al. (2022) (38,39). In short, to test these
175 assumptions of the prediction model and aspects of unidimensionality, Chi-square statistics are
176 calculated to evaluate the item and person fit, as well as the overall fit of the scale (38,39). The
177 principal component analysis of residuals (PCAR) provides additional information in relation to the
178 unidimensionality of the scale (38–40) with eigenvalues and percentage variance accounted for by
179 each principal component. Further analysis of paired *t*-tests provides evidence that the scale is
180 unidimensional, if less than 5% significant difference is found when comparing the two subtests
181 based on positive and negative loadings on the first principal component of the PCAR.

182 The person separation reliability (PSR) identifies the measurement’s precision and indicates whether
183 we can reliably separate high from low person ability at a group ($PSR \geq 0.70$) or individual level
184 ($PSR > 0.90$) (38,39,41). Targeting is identified by floor and ceiling effects (>15% considered as
185 problematic), as well as the person mean location relative to the item location, which is by default
186 positioned on the logit scale at 0 logits \pm 1 standard deviation (38,39). The scale is well targeted

187 when the person mean location is within 0.5 logits of the item mean (38,39). We identify local item
188 dependence through residual correlations (42). A correlation of at least 0.2 above the average residual
189 item correlation indicates that this pair of items have more in common with each other than with the
190 whole scale (42).

191 **3 Results**

192 We recruited 407 adults with (sub)acute stroke (63.2 ± 16.0 years of age; 157 women). The
193 demographic and clinical details are presented in **Table 1**. All patients were admitted to the
194 rehabilitation center (Neurocenter, Luzerner Kantonsspital, Lucerne, Switzerland) for inpatient
195 neurorehabilitation between January 2014 and November 2016 (24).

196 **3.1 Rasch-based UL-LIMOS**

197 The overall fit, item and person fit, PSR, floor-and ceiling percentages, and PCAR are shown in
198 **Table 2**. All items and persons fit the model. The UL-LIMOS fit the Rasch model without the need to
199 remove or rescore items. The individual item fit is displayed in **Table 3**. The threshold map for the
200 Rasch-converted UL-LIMOS is displayed in **Figure 1A**. This threshold map can be used in the
201 hospital or rehabilitation center for individual patient assessment. The person-item threshold
202 distribution is shown in **Figure 1B**. The total score of Rasch UL-LIMOS is displayed in **Table 4**,
203 with the conversion from the original ordinal scores (0 to 20 points) to logits, and logits further
204 converted to a 0 to 100 scoring.

205 Around 13.00% of participants obtained a maximum score, meaning there was a small but not
206 problematic ceiling effect. There was also a small (2.70%) floor effect. The average person mean
207 location was 1.32 ± 2.99 logits, indicating that the items were too easy for this group of adults with
208 (sub)acute stroke. PSR was 0.90, meaning the scale can reliably distinguish individuals of different
209 ability levels for decision-making in research and in the hospital or rehabilitation center (39).

210 The PCAR's eigenvalue on the first contrast was 2.46 with 49.23% explained variance on the first
211 principal component. Further analysis of paired *t*-tests revealed that 0.89% of person locations are
212 significantly different when comparing the two subtests formed based on positive (items 1 and 2) and
213 negative loadings (items 3 and 5) on the first principal component, thereby confirming the
214 unidimensionality of the scale.

215 Only one pair of items (items 1 "*lifting and carrying objects*" and item 2 "*fine hand use*"; $r = 0.65$)
216 was above 0.2 of the average residual item correlations ($r = 0.43$). Both specifically identify hand
217 use, which could explain why they are more strongly related to each other than to the rest of the
218 items.

219 **4 Discussion**

220 This study presents a new valid ICF-based observation scale (UL-LIMOS) to evaluate dependency
221 during upper limb use in daily life in adults with stroke at the ICF-participation level. Structural
222 validity of the UL-LIMOS was evaluated with Rasch analysis in 407 adults with (sub)acute stroke,
223 demonstrating that UL-LIMOS fit the model, without problematic floor or ceiling effects, and with a
224 high PSR of 0.90, which allows clinicians and researchers to use the scale for individual decision-
225 making.

226 Rasch analysis also provided insight into the hierarchy of difficulty of the five items. As expected,
227 based on the conceptual framework of upper limb movements, fine hand use requires more
228 dependency than arm and hand use. Tasks such as washing and dressing are even more difficult
229 because they require more awareness and interaction with the whole body, and/or require more
230 cognitive motor planning regarding the different motor action sequences to perform the activity. This
231 is reflected within the Rasch-based hierarchical order of the items.

232 Dressing appeared to be the most difficult item, confirming previous descriptions of dressing as a
233 complex skill that requires several physical motor function skills and cognitive abilities (43).
234 Notably, approximately 50% of adults with stroke still cannot dress independently six months post-
235 stroke (44). Cognitive deficits could be an important factor for this dependency, and, among the
236 spectrum of different cognitive factors, spatial neglect has been shown to have a major negative
237 impact on dressing skills in adults with a right hemispheric stroke (43,45). Upper limb apraxia has
238 been shown to affect dressing in adults with a left hemispheric stroke (38,39). In our sample, 78%
239 had deficits in cognition (30–32), and, more specifically, 39% had spatial neglect (33), and 35% had
240 moderate to severe apraxia (34–36), further confirming the problematic factors, influencing motor
241 actions that were previously identified in the literature. We confirm earlier findings in the literature
242 that more adults with right hemispheric stroke exhibit neglect and more adults with a left hemispheric
243 stroke have upper limb apraxia. However, the influence of these cognitive disorders on the use of the
244 upper limb in daily life needs to be studied, and assessed comprehensively, in much more detail in
245 future studies.

246 Evaluating upper limb motor impairment and/or activity in a structured, laboratory-based setting
247 ('capacity'), such as Fugl-Meyer (8), MESUPES (46), and ARAT (9), and comparing those results to
248 their performance level with UL-LIMOS, measuring reliance on others for upper limb use in daily
249 life, is important because these outcomes may not always line up. Adults with stroke may have the
250 ability to recruit motor units to perform specific motor actions in a laboratory setting but may not be
251 able to generate the necessary motor programs or have the necessary cognitive processing skills to
252 perform tasks in a more unstructured and more complex environment such as is the case with daily
253 life activities.

254 Evaluation scales are often used in the hospital and rehabilitation settings to provide some estimates
255 to patients regarding their recovery potential and which treatments would be most appropriate for
256 them. Therefore, early prediction algorithms have gained much attraction in recent years (47–51).
257 Yet, the current upper limb prediction models test upper limb motor function in a structured
258 laboratory setting,(47) which does not reflect actual upper limb use in daily life. In addition, these
259 studies often exclude adults with cognitive deficits post-stroke (48–50,52). Thus, the predictions are
260 only applicable to a subset of adults with stroke given the high prevalence of cognitive deficits after
261 stroke, including spatial neglect (53) and apraxia (35). Interestingly, Stinear et al. (2017), who
262 developed the prediction algorithms PREP and PREP-2, emphasized the importance of including
263 cognitive factors in future prediction paradigms, because these factors influenced upper limb
264 outcomes (48,51). This stresses the need for new predictive models that consider the evaluation of
265 dependency on others during upper limb use in daily life. The UL-LIMOS, which describes
266 dependency on others during spontaneous upper limb use in daily life, can also be used in patients
267 with neglect and apraxia. Therefore, UL-LIMOS, as well as measures of neglect and apraxia could
268 therefore be valuable factors in future predictive models of upper limb motor recovery after stroke.

269 Our study has limitations. Adults with stroke were recruited in only one neurorehabilitation center in
270 Switzerland, which could limit the generalizability to other countries with different cultures.

271 Furthermore, targeting could be improved in the future by adding more difficult items to the scale.
272 Other psychometrics such as sensitivity to change need to be performed on the 5-item UL-LIMOS.
273 Lastly, the scale also needs to be validated in adults with chronic stroke to ensure the generalizability
274 of the results.

275 In conclusion, we present a new 5-item Rasch-based UL-LIMOS scale, which was validated in 407
276 adults with acute or subacute stroke. We recommend validation of the UL-LIMOS in other countries
277 to account for cultural differences. The UL-LIMOS could also be validated in chronic stroke stages
278 when adults have returned to their home setting. A comparison of the UL-LIMOS data with self-
279 reported measurements or with accelerometers could potentially lead to changes to the existing core
280 datasets recommended for the evaluation of upper limb performance of adults with stroke (10,11).

281 **5 Conflict of Interest**

282 The authors declare that the research was conducted in the absence of any commercial or financial
283 relationships that could be construed as a potential conflict of interest.

284 **6 Author Contributions**

285 TV, BO, JV, and TN contributed to the conception and design of the study. TV organized the
286 database. AVDW performed the statistical analysis and the interpretation of the results, and she wrote
287 the first draft of the manuscript. TV wrote sections of the manuscript. All authors contributed to the
288 manuscript revision. They also read and approved the submitted version.

289 **7 Funding**

290 This study was supported by Swiss National Science Foundation grants (T.N. 320030_140696 and
291 320030_169789) and Innosuisse grant (T.V. 52272.1 IP-SBM).

292 **8 Acknowledgments**

293 We thank the neurorehabilitation team involved in the data collection and all participants for their
294 time and investment. We like to express our deep gratitude to Marc Noël for the critical review of the
295 manuscript.

296 **9 Data Availability Statement**

297 Original datasets requests can be directed to the first or last author.

298 **10 Contribution to the field statement**

299 Many adults with stroke experience problems using their affected arm in daily life. Until today,
300 however, there is no measure that evaluates how much adults with stroke need to rely on others when
301 using the affected arm in daily life.

302 To bridge this important gap, we have developed a new 5-item “Upper Limb Lucerne ICF-based
303 Multidisciplinary Observation Scale (UL-LIMOS)”, which reliably assesses dependency on others
304 during affected arm use in daily life in adults with stroke. Our study examining 407 adults with
305 (sub)acute stroke shows that questions in the UL-LIMOS are valid and are measuring what they are
306 supposed to measure. The UL-LIMOS could be used in new predictive models of upper limb motor

307 recovery after stroke because of the demonstrated structural validity without problematic floor or
308 ceiling effects. Our results also show that clinicians and researchers can use the scale for individual
309 decision-making.

310

311 **11 References**

- 312 1. Lawrence ES, Coshall C, Dundas R, Stewart J, Rudd AG, Howard R, et al. Estimates of the prevalence of
313 acute stroke impairments and disability in a multiethnic population. *Stroke*. 2001;32(6):1279–84.
- 314 2. Veerbeek JM, Kwakkel G, van Wegen EEH, Ket JCF, Heymans MW. Early prediction of outcome of
315 activities of daily living after stroke: a systematic review. *Stroke*. 2011 May;42(5):1482–8.
- 316 3. Chausson N, Olindo S, Cabre P, Saint-Vil M, Smadja D. Five-year outcome of a stroke cohort in
317 Martinique, French West Indies: Etude Realisee en Martinique et Centree sur l'Incidence des Accidents
318 vasculaires cerebraux *Stroke* [Internet]. 2010; Available from:
319 <https://www.ahajournals.org/doi/abs/10.1161/strokeaha.109.573402>
- 320 4. Schwab-Malek S, Vatankhah B, Bogdahn U, Horn M, Audebert HJ. Depressive symptoms and quality of
321 life after thrombolysis in stroke: the TEMPiS study. *J Neurol*. 2010 Nov;257(11):1848–54.
- 322 5. Aarnio K, Rodríguez-Pardo J, Siegerink B, Hardt J, Broman J, Tulkki L, et al. Return to work after
323 ischemic stroke in young adults [Internet]. Vol. 91, *Neurology*. 2018. p. e1909–17. Available from:
324 <http://dx.doi.org/10.1212/wnl.00000000000006510>
- 325 6. Alt Murphy M, Resteghini C, Feys P, Lamers I. An overview of systematic reviews on upper extremity
326 outcome measures after stroke. *BMC Neurol*. 2015 Mar 11;15:29.
- 327 7. World Health Organization. International Classification of Functioning, Disability and Health: ICF.
328 World Health Organization; 2001. 228 p.
- 329 8. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a
330 method for evaluation of physical performance. *Scand J Rehabil Med*. 1975;7(1):13–31.
- 331 9. Lyle RC. A performance test for assessment of upper limb function in physical rehabilitation treatment
332 and research. *Int J Rehabil Res*. 1981;4(4):483–92.
- 333 10. Kwakkel G, Lannin NA, Borschmann K, English C, Ali M, Churilov L, et al. Standardized Measurement
334 of Sensorimotor Recovery in Stroke Trials: Consensus-Based Core Recommendations from the Stroke
335 Recovery and Rehabilitation Roundtable. *Neurorehabil Neural Repair*. 2017 Sep;31(9):784–92.
- 336 11. Pohl J, Held JPO, Verheyden G, Murphy MA, Engelter S, Flöel A, et al. Corrigendum: Consensus-Based
337 Core Set of Outcome Measures for Clinical Motor Rehabilitation After Stroke—A Delphi Study
338 [Internet]. Vol. 12, *Frontiers in Neurology*. 2021. Available from:
339 <http://dx.doi.org/10.3389/fneur.2021.697935>
- 340 12. Mahoney FI, Barthel DW. Barthel Index [Internet]. *PsycTESTS Dataset*. 2013. Available from:
341 <http://dx.doi.org/10.1037/t02366-000>
- 342 13. Keith RA, Granger CV, Hamilton BB, Sherwin FS. The functional independence measure: a new tool for
343 rehabilitation. *Adv Clin Rehabil*. 1987;1:6–18.
- 344 14. Harvey RL. Predictors of Functional Outcome Following Stroke. *Phys Med Rehabil Clin N Am*. 2015
345 Nov;26(4):583–98.

- 346 15. Hsueh IP, Lee MM, Hsieh CL. Psychometric characteristics of the Barthel activities of daily living index
347 in stroke patients. *J Formos Med Assoc.* 2001 Aug;100(8):526–32.
- 348 16. Hsueh IP, Lin JH, Jeng JS, Hsieh CL. Comparison of the psychometric characteristics of the functional
349 independence measure, 5 item Barthel index, and 10 item Barthel index in patients with stroke. *J Neurol*
350 *Neurosurg Psychiatry.* 2002 Aug;73(2):188–90.
- 351 17. Duncan PW, Samsa GP, Weinberger M, Goldstein LB, Bonito A, Witter DM, et al. Health status of
352 individuals with mild stroke. *Stroke.* 1997 Apr;28(4):740–5.
- 353 18. Taub E, Miller NE, Novack TA, Cook EW 3rd, Fleming WC, Nepomuceno CS, et al. Technique to
354 improve chronic motor deficit after stroke. *Arch Phys Med Rehabil.* 1993 Apr;74(4):347–54.
- 355 19. Penta M, Tesio L, Arnould C, Zancan A, Thonnard JL. The ABILHAND questionnaire as a measure of
356 manual ability in chronic stroke patients: Rasch-based validation and relationship to upper limb
357 impairment. *Stroke.* 2001;32(7):1627–34.
- 358 20. Taub E, Crago JE, Uswatte G. Constraint-induced movement therapy: A new approach to treatment in
359 physical rehabilitation [Internet]. Vol. 43, *Rehabilitation Psychology.* 1998. p. 152–70. Available from:
360 <http://dx.doi.org/10.1037/0090-5550.43.2.152>
- 361 21. Doman CA, Waddell KJ, Bailey RR, Moore JL, Lang CE. Changes in Upper-Extremity Functional
362 Capacity and Daily Performance During Outpatient Occupational Therapy for People With Stroke. *Am J*
363 *Occup Ther.* 2016 May-Jun;70(3):7003290040p1–7003290040p11.
- 364 22. Lang CE, Waddell KJ, Barth J, Holleran CL, Strube MJ, Bland MD. Upper Limb Performance in Daily
365 Life Approaches Plateau Around Three to Six Weeks Post-stroke. *Neurorehabil Neural Repair.* 2021
366 Oct;35(10):903–14.
- 367 23. Ottiger B, Vanbellingen T, Gabriel C, Huberle E, Koenig-Bruhin M, Pflugshaupt T, et al. Correction:
368 Validation of the New Lucerne ICF Based Multidisciplinary Observation Scale (LIMOS) for Stroke
369 Patients. *PLoS One.* 2015 Jul 21;10(7):e0134186.
- 370 24. Van de Winckel A, Ottiger B, Bohlhalter S, Nyffeler T, Vanbellingen T. Comprehensive ADL Outcome
371 Measurement after Stroke: Rasch Validation of the Lucerne ICF-Based Multidisciplinary Observation
372 Scale (LIMOS) [Internet]. *Archives of Physical Medicine and Rehabilitation.* 2019. Available from:
373 <http://dx.doi.org/10.1016/j.apmr.2019.02.012>
- 374 25. Vanbellingen T, Ottiger B, Pflugshaupt T, Mehrholz J, Bohlhalter S, Nef T, et al. The Responsiveness of
375 the Lucerne ICF-Based Multidisciplinary Observation Scale: A Comparison with the Functional
376 Independence Measure and the Barthel Index. *Front Neurol* [Internet]. 2016;7. Available from:
377 <https://www.frontiersin.org/articles/10.3389/fneur.2016.00152>
- 378 26. Ottiger B, Vanbellingen T, Gabriel C, Huberle E, Koenig-Bruhin M, Pflugshaupt T, et al. Validation of
379 the new Lucerne ICF based Multidisciplinary Observation Scale (LIMOS) for stroke patients. *PLoS One.*
380 2015 Jun 25;10(6):e0130925.
- 381 27. Nyffeler T, Vanbellingen T, Kaufmann BC, Pflugshaupt T, Bauer D, Frey J, et al. Theta burst stimulation
382 in neglect after stroke: functional outcome and response variability origins [Internet]. Vol. 142, *Brain.*
383 2019. p. 992–1008. Available from: <http://dx.doi.org/10.1093/brain/awz029>
- 384 28. Bernhardt J, Hayward KS, Kwakkel G, Ward NS, Wolf SL, Borschmann K, et al. Agreed definitions and
385 a shared vision for new standards in stroke recovery research: The Stroke Recovery and Rehabilitation
386 Roundtable taskforce. *Int J Stroke.* 2017 Jul;12(5):444–50.

- 387 29. Vandenberg JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening
388 the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Int J*
389 *Surg*. 2014 Dec;12(12):1500–24.
- 390 30. Zuo L, Dong Y, Zhu R, Jin Z, Li Z, Wang Y, et al. Screening for cognitive impairment with the Montreal
391 Cognitive Assessment in Chinese patients with acute mild stroke and transient ischaemic attack: a
392 validation study. *BMJ Open*. 2016 Jul 12;6(7):e011310.
- 393 31. Feng Y, Zhang J, Zhou Y, Chen B, Yin Y. Concurrent validity of the short version of Montreal Cognitive
394 Assessment (MoCA) for patients with stroke [Internet]. Vol. 11, *Scientific Reports*. 2021. Available
395 from: <http://dx.doi.org/10.1038/s41598-021-86615-2>
- 396 32. Webb SS, Hobden G, Roberts R, Chiu EG, King S, Demeyere N. Validation of the UK English Oxford
397 cognitive screen-plus in sub-acute and chronic stroke survivors. *Eur Stroke J*. 2022 Dec;7(4):476–86.
- 398 33. Marques CLS, de Souza JT, Gonçalves MG, da Silva TR, da Costa RDM, Modolo GP, et al. Validation
399 of the Catherine Bergego Scale in patients with unilateral spatial neglect after stroke. *Dement*
400 *Neuropsychol*. 2019 Jan-Mar;13(1):82–8.
- 401 34. Vanbellingen T, Kersten B, Van de Winckel A, Bellion M, Baronti F, Müri R, et al. A new bedside test of
402 gestures in stroke: the apraxia screen of TULIA (AST). *J Neurol Neurosurg Psychiatry*. 2011
403 Apr;82(4):389–92.
- 404 35. Vanbellingen T, Kersten B, Van Hemelrijk B, Van de Winckel A, Bertschi M, Müri R, et al.
405 Comprehensive assessment of gesture production: a new test of upper limb apraxia (TULIA). *Eur J*
406 *Neurol*. 2010 Jan;17(1):59–66.
- 407 36. Van Hemelrijk B, Vanbellingen T, Van de Winckel A, De Weerd W, Bohlhalter S. A new test to
408 measure upper limb apraxia (TULIAS): A reliability study. *Mov Disord*. 2007;22:S213–4.
- 409 37. Vanbellingen T, Ottiger B, Maaijwee N, Pflugshaupt T, Bohlhalter S, Müri RM, et al. Spatial Neglect
410 Predicts Upper Limb Use in the Activities of Daily Living. *Cerebrovasc Dis*. 2017 Jun 13;44(3-4):122–7.
- 411 38. Mallinson T, Kozlowski AJ, Johnston MV, Weaver J, Terhorst L, Grampurohit N, et al. A Rasch
412 Reporting Guideline for Rehabilitation Research (RULER): The RULER Statement. *Arch Phys Med*
413 *Rehabil* [Internet]. 2022; Available from: <http://dx.doi.org/10.1016/j.apmr.2022.03.013>
- 414 39. Van de Winckel A, Kozlowski AJ, Johnston MV, Weaver J, Grampurohit N, Terhorst L, et al. Reporting
415 Guideline for RULER: Rasch Reporting Guideline for Rehabilitation Research – Explanation &
416 Elaboration manuscript. *Arch Phys Med Rehabil* [Internet]. 2022; Available from:
417 <http://dx.doi.org/10.1016/j.apmr.2022.03.019>
- 418 40. Linacre JM. Dimensionality: contrasts & variances. A user’s guide to Winsteps Ministep Rasch-model
419 computer programs (version 3 81 0) Retrieved from [http://www.winsteps](http://www.winsteps.com/winman/principalcomponents.htm)
420 [com/winman/principalcomponents htm](http://www.winsteps.com/winman/principalcomponents.htm) [Internet]. 2014; Available from:
421 <https://www.winsteps.com/winman/principalcomponents.htm>
- 422 41. Reliability and separation of measures [Internet]. [cited 2021 Jan 30]. Available from:
423 <https://www.winsteps.com/winman/reliability.htm>
- 424 42. Christensen KB, Makransky G, Horton M. Critical Values for Yen’s Q_3 : Identification of Local
425 Dependence in the Rasch Model Using Residual Correlations [Internet]. Vol. 41, *Applied Psychological*
426 *Measurement*. 2017. p. 178–94. Available from: <http://dx.doi.org/10.1177/0146621616677520>
- 427 43. Walker MF, Lincoln NB. Factors influencing dressing performance after stroke. *J Neurol Neurosurg*

- 428 Psychiatry. 1991 Aug;54(8):699–701.
- 429 44. Edmans JA, Lincoln NB. The Relation between Perceptual Deficits after Stroke and Independence in
430 Activities of Daily Living. *Br J Occup Ther.* 1990 Apr 1;53(4):139–42.
- 431 45. Walker MF, Sunderland A, Fletcher-Smith J, Drummond A, Logan P, Edmans JA, et al. The DRESS
432 trial: a feasibility randomized controlled trial of a neuropsychological approach to dressing therapy for
433 stroke inpatients. *Clin Rehabil.* 2012 Aug;26(8):675–85.
- 434 46. Van de Winckel A, Feys H, van der Knaap S, Messerli R, Baronti F, Lehmann R, et al. Can quality of
435 movement be measured? Rasch analysis and inter-rater reliability of the Motor Evaluation Scale for
436 Upper Extremity in Stroke Patients (MESUPES). *Clin Rehabil.* 2006 Oct;20(10):871–84.
- 437 47. Nijland RHM, van Wegen EEH, Harmeling-van der Wel BC, Kwakkel G, EPOS Investigators. Presence
438 of finger extension and shoulder abduction within 72 hours after stroke predicts functional recovery: early
439 prediction of functional outcome after stroke: the EPOS cohort study. *Stroke.* 2010 Apr;41(4):745–50.
- 440 48. Stinear CM, Barber PA, Petoe M, Anwar S, Byblow WD. The PREP algorithm predicts potential for
441 upper limb recovery after stroke. *Brain* 135 (Pt 8), 2527--2535. 2012.
- 442 49. van der Vliet R, Selles RW, Andrinopoulou ER, Nijland R, Ribbers GM, Frens MA, et al. Predicting
443 Upper Limb Motor Impairment Recovery after Stroke: A Mixture Model. *Ann Neurol.* 2020
444 Mar;87(3):383–93.
- 445 50. Selles RW, Andrinopoulou ER, Nijland RH, van der Vliet R, Slaman J, van Wegen EE, et al.
446 Computerised patient-specific prediction of the recovery profile of upper limb capacity within stroke
447 services: the next step. *J Neurol Neurosurg Psychiatry.* 2021 Jan 21;92(6):574–81.
- 448 51. Stinear CM, Byblow WD, Ackerley SJ, Smith MC, Borges VM, Barber PA. PREP2: A biomarker-based
449 algorithm for predicting upper limb function after stroke. *Ann Clin Transl Neurol.* 2017 Nov;4(11):811–
450 20.
- 451 52. Krakauer JW, Marshall RS. The proportional recovery rule for stroke revisited. *Ann Neurol.* 2015
452 Dec;78(6):845–7.
- 453 53. Azouvi P, Samuel C, Louis-Dreyfus A, Bernati T, Bartolomeo P, Beis JM, et al. Sensitivity of clinical
454 and behavioural tests of spatial neglect after right hemisphere stroke. *J Neurol Neurosurg Psychiatry.*
455 2002 Aug;73(2):160–6.

456

457

458

459

460 **12 Figure Legends**

461 **Figure 1**

462 **A. The Rasch-based UL-LIMOS scale: Item threshold map:** The item threshold map depicts
463 the difficulty of the items from the easiest item at the top to the most difficult item at the
464 bottom along with the scoring categories. These item thresholds are matched on the same
465 logit scale (horizontal black line at the bottom of the picture) as the person's ability. This is a
466 visual depiction of the interval scale, using the same color coding for each item threshold.
467 This demonstrates that with increasing ability, it is easier to get a higher score on an easy item
468 than on a difficult item. It also shows what score would be expected for each item, based on a
469 person's ability.

470 **B. The Rasch-based UL-LIMOS scale: Person-item threshold distribution:** The ability of
471 the persons (top, pink bars) is plotted on the same logit scale as the difficulty of the item
472 thresholds (bottom, blue bars). The histograms depict the frequency of persons at a certain
473 ability level, from a low ability on the left to a high ability on the right side of the ruler. The
474 number of item thresholds is organized in increasing difficulty levels from the easiest on the
475 left to the most difficult item thresholds on the right side of the ruler.

476

477

Table 1. Demographic and clinical characteristics of adults with (sub)acute stroke

Demographic and clinical characteristics	n = 407
Age , years, median (IQR)	63 (53-78)
Sex , n (m/f)	157/250
Ischemic , n (%)	276 (68%)
Hemorrhagic , n (%)	131 (32%)
Stroke laterality , n	
Right , n (%)	149 (37%)
Left , n (%)	193 (47%)
Both , n (%)	65 (16%)
Time post stroke , days, median (IQR)	13 (8-42)
Affected arterial territory , n	
arteria cerebri anterior , n (%)	21 (5%)
arteria cerebri media , n (%)	225 (55%)
arteria cerebri posterior , n (%)	83 (20%)
Other	78 (19%)
Cognitive function (MoCA) , mean \pm SD	19.14 \pm 6.63
Unilateral neglect (CBS) , mean \pm SD	3.78 \pm 6.55
Apraxia (AST) , mean \pm SD	9.01 \pm 3.11

Legend: AST = Apraxia Screen of the test of upper limb apraxia (TULIA); CBS =

Catherine Bergego Scale; MoCA = Montreal Cognitive Assessment; n = number of patients

478

Table 2. The overall fit, item and person fit, and person separation reliability (PSR), floor-and ceiling values and percentages, and principal component analysis of residuals (PCAR)

Analysis	Items	Rating scale categories	Person mean (SD) logits	Mean error variance	Floor effect n (%)	Ceiling effect n (%)	Overall Chi-square (DF) p-value	PSR	Items with disordered thresholds (n)	Misfitting items (n)	PCAR Eigenvalue 1 st contrast (%)	Misfitting persons n (%)
All Items admission data (n=403, retrospective sample)	5	25	1.32 (2.99)	0.90	11 (2.70)	53 (13.0)	28.91 (15) p=0.016	0.90	0	0	2.46 (49.23)	0 (0.00)

Table 3. Item fit statistics of the Rasch-based UL-LIMOS scale

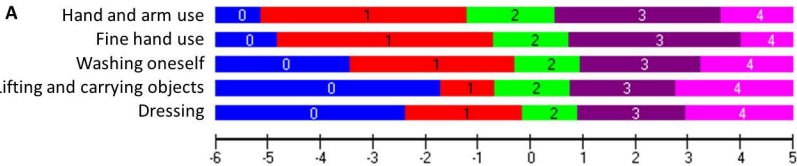
Item number	Location (logits)	SE	Chi Square	Item Fit Residual	<i>p</i> -value
UL-LIMOS					
1. Hand and arm use	-0.56	0.09	3.33	-1.60	0.34
2. Fine hand use	-0.19	0.09	2.24	-0.97	0.52
3. Washing oneself	0.12	0.09	8.95	-2.05	0.03
4. Lifting and carrying objects	0.29	0.08	6.11	6.32	0.10
5. Dressing	0.33	0.08	8.07	-2.82	0.04

Legend: SE: Standard Error; *p*-values are Bonferroni corrected for multiple comparisons ($\alpha=0.01$ for 5 items)

Table 4. Converted scoring system for the total UL-LIMOS score.

Raw score	logits	Converted logits to 0-100
UL-LIMOS		
0	-6.38	0
1	-5.17	10
2	-4.09	19
3	-3.21	27
4	-2.37	34
5	-1.79	39
6	-1.33	43
7	-0.92	46
8	-0.56	49
9	-0.22	52
10	0.11	55
11	0.44	58
12	0.79	61
13	1.16	64
14	1.57	67
15	2.04	71
16	2.56	76
17	3.12	80

18	3.74	86
19	4.50	92
20	5.45	100



B

