

1 Reliability and validity of physical function tests and ADL survey
2 questions in women living in rural, highland Ethiopia

3
4 Jenna Golan^{1*}, Anna Thalacker-Mercer², John Hoddinott¹

5 ¹Division of Nutritional Sciences, Cornell University, Ithaca, New York, USA

6
7 ²Department of Cell, Development and Integrative Biology, University of Alabama at
8 Birmingham, Birmingham, AL, USA

9 * Corresponding author

10 jmg559@cornell.edu

11

12

13 **Abstract**

14 In rural, highland Ethiopia, physical function, which is the physical ability to fulfill one's daily
15 roles and responsibilities, may be compromised by a lack of access to nutrition, healthcare, and
16 sanitation. Decreased physical function would be detrimental to health and income-generating
17 activities. Unfortunately, there is a lack of validated methods to measure physical function in
18 adult women in this region. This validity study evaluated the feasibility and reliability of
19 physical tests, including the sit-to-stand (STS) and usual gait speed (UGS). The physical tests
20 were used to assess the validity of context-appropriate activities of daily living (ADL) questions.
21 The results of the STS were used to calculate a power index (P_{sts}) which accounted for body
22 mass and leg length. Feasibility was ascertained qualitatively. Reliability was assessed by
23 comparing the results of the tests and questions between each visit using either Cohen's κ or
24 Pearson's ρ . Validity was assessed by regressing the responses to the ADL questions against
25 P_{sts} , controlling for relevant participant characteristics. Study participants consisted of 316
26 women between the age of 18 and 45 years, living in rural Tigray, Ethiopia, that had previously
27 participated in an impact evaluation of a safety net program. Over a one-week period,
28 participants completed the STS and UGS tests and responded to the ADL questionnaires three
29 times. P_{sts} was determined to be a feasible, reliable, and valid physical function test for women
30 in rural, highland Ethiopia. UGS lacked feasibility and reliability. The validity of the ADLs was
31 inconclusive. The P_{sts} will be an essential tool for improving physical function and,
32 subsequently, health and quality of life in rural Ethiopia.

33 **Introduction**

34 A person's ability to functionally, independently, and physically fulfill their activities of
35 daily living (ADL) is known as physical function (1,2). It is both a cause and consequence of
36 body size and composition and physical activity. Intrinsicly, physical function is essential for
37 overall health and quality of life (3,4). Instrumentally, physical function can identify people at
38 risk for further declines in health (5–9), requiring additional care (4), or having future health
39 expenditures (10). Additionally, studies have used physical function as an outcome measure to
40 determine the impact of interventions targeting health (11–13). Outside of health measures,
41 physical function impacts social networks (14,15) and income-generating activities (16–18).

42 Medical practitioners and researchers in high-income countries have extensively
43 measured physical function in a variety of populations, such as older adults with and without
44 dementia (5,19–23) and people with chronic illness or pain (e.g., arthritis) (24–26). However,
45 measuring physical function has been underused in low- and middle-income countries (LMIC)
46 populations, with few exceptions (19,27–29).

47 Measuring physical function in LMIC is essential because daily life is physically
48 demanding. According to the 2016 Ethiopia Demographic and Health Survey (DHS), of the
49 employed women living in rural Ethiopia, approximately 65% were employed in skilled manual
50 labor, unskilled manual labor, or agriculture. Additionally, 53% of women in rural households
51 spent more than 30 minutes fetching drinking water daily (30). Furthermore, populations living
52 in rural Ethiopia are exposed to many risk factors that would negatively impact their physical
53 function, including access to nutrition (31–33), healthcare, and sanitation (34–36). According to
54 the 2016 DHS report, 25% of women living in rural Ethiopia were underweight (BMI < 18.5

55 kg/m^2), and 25% were anemic (30). Declines in physical function would significantly impact both
56 necessary income-generating activities and activities needed for daily life.

57 Physical function is measured through physical tests, or validated survey questions
58 focused on the respondents' ability to complete tasks required for personal care and independent
59 living (37). Several tests, including usual gait speed (UGS) and sit-to-stand tests (STS), do not
60 require specialized training or resources, making them a viable option for measuring physical
61 function in resource-limited settings. Gait speed strongly relates to pathophysiological
62 mechanisms (38) and is considered a global marker of well-being (39). The STS requires a
63 combination of muscle power and balance (40). It uses upper and lower-body musculature,
64 requires continuous balance adjustments, and employs a motor pattern commonly used during
65 ADLs (41). Both tests collect continuous outcome measures, which collect data on a gradient at
66 the upper end of the functional spectrum, making them more useful in young and active
67 populations (5,42,43).

68 Survey questions require less time and training to administer than direct tests, making
69 them suitable for population studies. ADL surveys are one example of the validated survey
70 questions used to measure physical function (44). They use population-specific questions to
71 determine how well participants complete their ADLs (19,45). For example, surveys for elderly
72 or infirmed populations often ask about participants' ability to bathe, dress, and feed themselves
73 (1,20,37,44).

74 Physical function is not well understood in rural Ethiopia because the physical function
75 assessment tools have not been studied in this setting. The objectives of this study were to
76 determine the feasibility, reliability, and validity of the UGS, STS, and ADLs survey to measure
77 physical function in women living in rural highland Ethiopia. By understanding physical

78 function in the region, future research could identify those needing additional healthcare and
79 factors influencing health.

80 **Materials and Methods**

81 **Study setting**

82 Data were collected in the mountainous Central, Eastern, and South Zones of Tigray,
83 Ethiopia. Most people in the rural highland regions have livelihoods based on rainfed
84 agriculture. A short rainy season occurs between March and May, and the main rainy season
85 occurs between June and October. During the main rainy season, more than 90 percent of the
86 country's total crop production occurs, with harvesting occurring between October and
87 December (46). In addition, during the small rainy season, the poorest households in the regions
88 receive additional support from a government social safety net called the Productive Safety Net
89 Program (PSNP). The PSNP provides cash or food payments for work on community projects
90 such as terracing land, building roads, water reclamation, and reforestation efforts (46,47).

91 Participants in this study belonged to households who had previously taken part in an
92 evaluation of the PSNP (46). According to this evaluation, the villages included in this study
93 were remote. The average travel time to the market was 85 minutes on foot. Only 1% of
94 households had water piped into their dwelling; the average time to fetch water and return was
95 45 minutes. Female respondents spent most of their time conducting childcare, cooking/ eating,
96 domestic activities, and personal care during the dry season. During the main rainy season, they
97 spent most of their time on childcare, cooking/ eating, and food production activities (46).

98 **Study subjects**

99 Participants in this study resided in households that had previously taken part in an
100 evaluation of the PSNP (46). They were recruited based on the zones, counties, and villages
101 where they lived. Within the zones, counties were excluded if the PSNP was inactive. Within
102 each county, villages were excluded if they did not meet the criterion for a concurrently running
103 study on physical exertion (48) which required two households that received benefits from the
104 public works program and two that did not. All PSNP evaluation participants living in the
105 selected villages were screened for eligibility. Prior studies that validated ADL questionnaires
106 ranged in sample size from 30 (49,50) to over 400 people (51). Inclusion criteria included
107 women aged 18-45 years who were not pregnant and were not planning on leaving the area
108 during data collection. One hundred and seventy-eight women participated in the first round of
109 data collection during the dry season, and 138 women participated in the second round of data
110 collection during the main rainy season, the agricultural round. One hundred and twenty-seven
111 women participated in data collection during both rounds; 51 only participated during the dry
112 season round and 11 during the agricultural round.

113 **Data collection**

114 Data collection took place over two four-week periods. The first four-week period,
115 April-May 2019, coincided with the small rainy season when minimal agricultural activities
116 occurred, and the PSNP was operational. The second, four-week period, September-October
117 2019, occurred during the main rainy season when households were engaged in farming
118 activities such as preparing the land, planting, and weeding. The PSNP was non-operational.

119 Over eight days, participants were visited up to six times. Participants answered survey
120 questions during three visits, including the ADL questions, and completed the physical function

121 tests. During the first visit, relevant anthropometric information, including weight (Seca, Model
122 874 dr, Hamburg, Germany) and leg length, were also measured. Women wore light clothing
123 while being weighed. Leg length was measured from the greater trochanter to the lateral
124 malleolus using a standard measuring tape. All anthropometric measurements were measured
125 twice, and the average of the two was used for analysis. Additional anthropometric,
126 demographic, and village information including was collected during the PSNP evaluation (46).

127 **Physical function instruments**

128 **Sit-to-stand**

129 A repeated STS test measures the time it takes a participant to rise from a seated position
130 without using their hands. This task requires a combination of muscle power and balance (40).
131 It uses upper and lower-body musculature, requires continuous balance adjustments, and
132 employs a motor pattern commonly used in ADLs (41). There is no standardized protocol for
133 this test, and several variations exist. For this study, we adapted the protocol used by Takai et al.
134 (52), modifying the seated position and the number of tests to make them appropriate for this
135 setting. The subjects started in a standing position, then sank into a low seated squat before
136 standing again. They squatted seven times as quickly as possible. The bottoms of their feet
137 remained in contact with the ground throughout the test. The low seated squat is a common
138 position for adults in developing countries to adopt when sitting or resting. Participants crossed
139 their arms at the wrists and held them against their chest to ensure they were unaided in rising.

140 A stopwatch recorded the time to the nearest 10th of a second. The STS test was
141 performed twice with an interval of 1 minute between the trials. The shorter of the two times
142 was used for data analysis. A power index (P_{sts}) was calculated using Equation 1. Previous

143 studies have found that the power index is correlated with the maximum voluntary isometric
144 knee extension force (52).

145

$$146 \quad P_{\text{sts}} = \ln \frac{L \times \text{body mass} \times g \times 7}{T_{\text{sts}}} \quad (1)$$

147

148 Where: P_{sts} is the power index from the test

149 L is the leg length measured from the greater trochanter to the lateral malleolus

150 Body mass is measured in kg to the nearest tenth

151 G is the acceleration of gravity (9.8 m/s²)

152 T_{sts} is the time it took to complete the test

153 **Usual gait speed**

154 During the UGS, participants were timed as they walked 20 meters at their usual pace.

155 Enumerators laid out pre-measured lengths of rope on a flat surface, free of obstacles, on stable
156 ground. The participants started three steps behind the start of the distance over which they were
157 recorded walking and continued walking until they were three steps beyond the finishing point.

158 The test was repeated three times, and the two closest times (out of three) were averaged and
159 used for data analysis.

160 **Activities of daily living survey**

161 ADL surveys have been used to assess physical function since they were developed in
162 1963 (53). Multiple versions of the survey exist. Each version asks the study participants if they
163 experienced difficulty completing routine activities such as getting in/out of bed, showering, and
164 preparing food. The routine activities selected for this version of an ADL survey were based on
165 pilot data from a “24-hour recall of time use” and “perceived energy exertion” survey conducted

166 by the study team. Three ADLs were selected that almost all participants had engaged in and
167 were not expected to vary much in difficulty throughout the year. The ADLs survey questions
168 collected information on participants' perceived difficulty (i) traveling to and from the market
169 (ADL - Market), (ii) cleaning and housework (ADL - Housework), and (iii) preparing food
170 (ADL - Food) due to the physical nature of the activity.

171 The participants ranked the level of difficulty encountered on an ordinal scale 1)
172 experienced no difficulty, 2) experienced some difficulty, 3) needed assistance to complete the
173 task, 4) avoided completing the task unless necessary, and 5) not responsible for the task. The
174 responses to the ADL survey questions were collapsed into binary values whether the participant
175 experienced no difficulty or any level of difficulty. Participants who were not responsible for a
176 task were marked as "missing." Additionally, a binary variable indicated whether a participant
177 responded positively to experiencing difficulties with any ADLs (ADL – Any), and another
178 variable captured the number of ADLs a participant reported difficulties (ADL – Number).

179 **Analysis**

180 Feasibility was assessed based on the training and tools needed to administer the tests and
181 qualitative reports from the enumerators on their ability to administer them. The reliability of
182 each physical function test was evaluated using the Pearson correlation coefficient (ρ). The
183 physical function tests from each visit were compared to the other visits with data. A coefficient
184 between 0 to 0.20 was poor agreement, 0.21 to 0.40 was fair agreement, 0.41 to 0.60 was
185 moderate/ acceptable agreement, 0.61 to 0.80 was substantial agreement, and 0.81 to 1.0 was
186 near perfect agreement. These classifications have been used in other validation studies (54).
187 Reliability for each ADL question was evaluated using Cohen's kappa statistic (κ). The same
188 cutoffs were used to assess agreement for the correlation coefficient (55).

189 Due to the low agreement between UGS and P_{sts} and low agreement of UGS between
190 visits, only P_{sts} was used to assess the validity of the ADL survey questions. One visit per
191 participant per round was randomly selected for analysis for the validity of the ADL survey.
192 First, a univariate analysis was conducted to determine if the responses to the ADL questions
193 were significantly associated with P_{sts} . Then, the variables were regressed with the rounds
194 collapsed and by round. In rounds where the p-value for an ADL question was less than 0.20,
195 the ADL question was fit into a multivariate regress with P_{sts} as the dependent variable, and
196 significant participant characteristics from that round were considered in the model. Significant
197 participants characteristics included BMI, underweight (BMI < 18.5 kg/m²), having a partner
198 live in the household, being food insecure as measured by food gap (46), self-reported short- or
199 long-term illness, and residing in a household that benefits from the PSNP. The best fit modeled
200 was determined using backward-stepwise regression by Bayesian information criterion (BIC).

201 **Ethical approval**

202 This study received ethical approval from the IRB at Cornell University (1902008596).

203 **Results**

204 Between the two rounds of data collection, participants were comparable in age,
205 education, BMI, marital status, PSNP beneficiary status, and their self-reported long-term and
206 short-term illnesses and injuries, **Table 1**. The mean age of the participants was 30 years, and
207 the range was 18 to 45. The average BMI was 19.4 kg/m². However, more participants were
208 underweight in the dry season round than in the agricultural (35.4% versus 31.6%). In addition,
209 there was a seasonal difference in food security, with more participants reporting a food gap in
210 the past six months during the dry season round (33.1%) than in the agricultural round (26.8).

211

Table 1: Study participant characteristics

Characteristic	Dry season round	Agricultural round	P-value
N	178	138	
Age (years)	30 (6.6)	30 (6.5)	0.46
Any education	50.0%	51.0%	0.90
BMI (kg/m^2)	19.4 (2.1)	19.5 (2.3)	0.54
Underweight (BMI < 18.5 kg/m^2)	35.4%	31.6%	0.48
Partner lives in household	85.4%	83.3%	0.62
Food insecure during the past 6 months	33.1%	26.8%	0.23
Long-term illness or injury	15.7%	14.5%	0.76
Short-term illness or injury	24.7%	26.8%	0.67
Benefits from PSNP	47.2%	44.9%	0.69
P_{sts}	9.9 (0.3)	9.9 (0.3)	0.89
UGS	8.7 (1.3)	8.5 (1.4)	0.26
ADL - Market	14.0%	21.0%	0.10
ADL - Housework	12.4%	20.3%	0.06
ADL - Food	12.4%	21.0%	0.04*
ADL - Any	17.4%	26.1%	0.06
ADL - Number			
Zero	82.6%	73.9%	0.06
One	5.1%	6.5%	0.58
Two	3.4%	2.9%	0.81
Three	9.0%	16.7%	0.04*

Values are presented as means with standard deviation in parentheses or as a percent of the study population. * Indicates that it is significant at $\alpha = 0.05$.

212

213 The average values of the two physical function tests were similar between rounds. More
214 participants reported difficulty with ADL - Any and each ADL during the agricultural round
215 compared to the dry season round. There was only a statistically significant difference (p-value
216 = 0.04) between those that reported difficulty preparing food for each round (12.4% vs 21.0%).
217 During the dry season round, 82.6% of participants reported experiencing no difficulties with
218 ADLs; this decreased to 73.9% in the agricultural round. In the dry season round, 9.0% of
219 participants reported difficulties with all three ADLs. This increased to 16.7% of participants in
220 the agricultural round (p-value = 0.04).

221 The enumerators reported that the STS was feasible to administer to the study population.
222 Participants could complete the number of repetitions but reported that it was challenging and
223 had started to fatigue. There were more reported difficulties in administering the UGS,
224 particularly in finding a sufficient area to administer the tests. The UGS was not feasible in
225 some households, making it an unsuitable test for use in rural Ethiopia.

226 The P_{sts} had a near-perfect agreement for each visit within a round, with a correlation
227 coefficient greater than 0.80 (**Table 2**), except for visit 1 compared to visit 3 in the dry season
228 round. There was substantial agreement between the values. Overall, UGS had a lower
229 agreement, with most comparisons having moderate agreement. There was substantial
230 agreement between visits 2 and 3 in the agricultural round (correlation coefficient = 0.62). There
231 was nearly always poor agreement between the UGS and P_{sts} between each visit for both rounds.

232

Table 2: Reliability of physical function tests

Test	Visit numbers	Dry season round		Agricultural round	
		N	Pearson's ρ	N	Pearson's ρ
P_{sts}					
	1 vs 2	176	0.81****	126	0.88****
	1 vs 3	147	0.78***	122	0.82****
	2 vs 3	147	0.81****	129	0.84****
UGS					
	1 vs 2	176	0.54**	124	0.59**
	1 vs 3	151	0.40*	117	0.40*
	2 vs 3	152	0.47**	121	0.62**
P_{sts} vs UGS					
	Visit 1	175	-0.22*	127	-0.12
	Visit 2	176	-0.10	132	-0.05
	Visit 3	147	-0.18	124	-0.05

* Indicates fair agreement, ** indicates moderate/ acceptable agreement, *** indicates substantial agreement, **** indicated near-perfect agreement.

233

234 The responses to each ADL question were compared between visits in each round, **Table**

235 **3.** During the dry season round, there was a near-perfect agreement in the responses to the ADL

236 questions between the visits (κ ranged from 0.78 to 0.97). During the agricultural round, there
 237 was either substantial or near-perfect agreement (κ ranged from 0.70 to 0.88).
 238

Table 3: Correlation of ADL questions

ADL question	Visit numbers	Dry season round		Agricultural round	
		N	Cohen's κ	N	Cohen's κ
ADL - Market					
	1 vs 2	174	0.89****	128	0.81****
	1 vs 3	150	0.81****	124	0.83****
	2 vs 3	150	0.89****	131	0.80****
ADL - Housework					
	1 vs 2	175	0.97****	128	0.77****
	1 vs 3	149	0.85****	124	0.70****
	2 vs 3	149	0.88****	131	0.84****
ADL - Food					
	1 vs 2	176	0.97****	128	0.84****
	1 vs 3	150	0.85****	124	0.86****
	2 vs 3	150	0.88****	131	0.88****
ADL - Any					
	1 vs 2	176	0.91****	128	0.76****
	1 vs 3	151	0.78****	124	0.70****
	2 vs 3	151	0.84****	131	0.71****
ADL - Number					
	1 vs 2	177	0.91****	137	0.74****
	1 vs 3	177	0.75****	137	0.65****

2 vs 3 177 0.79*** 137 0.71***

* Indicates fair agreement, ** indicates moderate/ acceptable agreement, *** indicates substantial agreement, **** indicated near-perfect agreement.

239

240 **Table 4** shows the association between P_{sts} and participant characteristics and ADL
 241 responses with the rounds combine and by round. Age, BMI, underweight, and self-reported
 242 long-term illness or injury were consistently significantly associated with P_{sts}. (p-value < 0.05).
 243 For every year increase in age, physical function as measured by the P_{sts} decreased. An increase
 244 in BMI was associated with improvements in physical function as measured by P_{sts}, while those
 245 underweight had lower physical function as measured by P_{sts}. Those that self-reported a long-
 246 term illness or injury had a lower physical function than those that did not. Among the ADL
 247 responses, none were significantly associated with P_{sts} when the rounds were combined. During
 248 the dry season round, ADL – Housework, ADL – Food, ADL – Any, and ADL – Number were
 249 significant at the $\alpha = 0.10$. Only the ADL - Market was significant at the $\alpha = 0.10$ level during
 250 the agricultural round.

251

Table 4: Association of P_{sts} with participant characteristics and ADL responses

Model parameter	All		Dry season round		Agricultural round	
	Coefficien t	p-value	Coefficien t	p-value	Coefficien t	p-value
N	316		178		138	
Age	-0.01	0.01**	-0.01	0.06*	-0.01	0.04**
Any education	-0.02	0.46	-0.01	0.88	-0.04	0.30
BMI (kg/m ²)	0.04	<0.01***	0.05	<0.01***	0.03	<0.01***

Underweight (BMI < 18.5 kg/m ²)	-0.15	<0.01***	-0.18	<0.01***	-0.11	0.02**
Partner lives in household	0.05	0.29	0.05	0.41	0.04	0.50
Food insecure	-0.05	0.13*	-0.01	0.78	-0.11	0.02**
Long term illness	-0.13	0.04**	-0.13	0.04**	-0.14	0.02**
Short term illness	0.01	0.78	0.02	0.74	-0.04	0.37
Safety net beneficiary	0.04	0.17*	0.03	0.51	0.06	0.16*
ADL – Market	-0.01	0.82	0.08	0.21	-0.10	0.07*
ADL – Housework	0.02	0.59	0.12	0.09*	-0.06	0.29
ADL – Food	0.03	0.52	0.12	0.08*	-0.05	0.33
ADL - Any	0.03	0.53	0.12	0.08*	-0.05	0.33
Number of ADLs	0.01	0.74	0.04	0.09*	-0.03	0.17

The mean value of $\ln(P_{sts})$ for each round was 9.92 with a standard deviation of 0.30 for round 1 and 0.25 for round 2. * Indicates statistical significance less than 0.20, qualifying for consideration in the multivariate model, ** Indicates statistical significance at <0.05, *** Indicates statistical significance <0.01

252

253 Multivariate models were constructed for each ADL question where the response was
 254 significant at $\alpha = 0.10$, Table 5. In the dry season round, the ADL - Housework and ADL - Any
 255 were significant after controlling for BMI. The ADL - Food and ADL – Number were
 256 significant after controlling for BMI and self-reported long-term illness or injury. For the
 257 agricultural round, ADL-Market was significant after controlling for BMI. In the dry season
 258 round, all significant ADL variables had a positive coefficient, indicating that they were
 259 positively associated with P_{sts} and had improved physical function compared to those that

260 responded that they had no difficulties. During the agricultural round, ADL - Market had a
 261 negative coefficient, indicating that those who reported difficulty had lower P_{sts} and lowered
 262 physical function than those who responded that they had no difficulties.
 263

Table 5: Multivariate model ADLs and P_{sts}

ADL question	ADL		BMI		Long-term illness or injury	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Dry season round						
Clean	0.13	0.04**	0.05	0.02**		
Food	0.15	0.02**	0.05	<0.01***	-0.12	0.04**
Any ADL	0.12	0.03**	0.05	<0.01***		
Number of ADL	0.05	0.03**	0.05	<0.01***	-0.12	0.05**
Agricultural round						
Travel	-0.10	0.04**	0.03	<0.01***		

The mean value of P_{sts} for each round was 9.92 with a standard deviation of 0.30 for round 1 and 0.25 for round 2. * Indicates statistical significance less than 0.20, ** Indicates statistical significance at <0.05, *** Indicates statistical significance <0.01

264

265

266 Discussion

267 This research found that P_{sts} (power index) is a feasible, reliable, and valid method to
 268 measure physical function in adult women living in rural highland Ethiopia. Measuring P_{sts}
 269 required very little training or materials, making it a low-cost, feasible option for research

270 conducted in resource-constrained settings. Additionally, enumerators reported that the test was
271 easy to administer and acceptable to the study participants. P_{sts} had high reliability with a near-
272 perfect agreement between visits in each round. Additionally, P_{sts} was associated with age, BMI,
273 and self-reported long-term illness or injury, indicating it is a valid measure of physical function.
274 P_{sts} is a promising tool for measuring physical function in similar populations. UGS was neither
275 feasible nor reliable, and it is not likely a good measure of physical function in adult women
276 living in rural highland Ethiopia. The ADL survey was reliable, but it had inconclusive validity.
277 The high reliability of the subjective scoring for each ADL among the rounds indicates that the
278 participants understood the question and interpreted it consistently. More work is needed for it
279 to be used to measure physical function in rural Ethiopia.

280 Strengths of this research include the careful attention to the development of the
281 methodologies and the administration of the STS and UGS in a novel setting. Using P_{sts} as a
282 measure of physical function is a strength of this paper. Previous work on physical function in
283 LMICs used grip strength (27,29,56). To the authors' knowledge, P_{sts} has not previously been
284 used in rural regions of LMIC. P_{sts} is a better measure of physical function than grip strength
285 because it requires muscle strength, mobility, balance, and coordination (40). In addition, P_{sts} is
286 more likely to reflect the necessary activities to fulfill a study participant's daily roles and
287 responsibilities (41).

288 There were limitations in this paper, particularly around the development of the ADLs.
289 The three activities were selected because they appeared in most responses in the time-use pilot
290 survey. There was minimal seasonal variation in the frequency or difficulty of completing these
291 activities. Unfortunately, logistical and time constraints prevented the piloting of these questions
292 before the study. Questions added to the agricultural round of data collection found that these

293 activities were the ones that were most likely to be altered when a participant was unwell.
294 However, this was not reflected in that data. There is a lack of information to help us understand
295 why the ADL questionnaire had limited validity or to aid in creating a new survey.

296 More research is needed to explore survey questions and activities that accurately
297 characterize this population's physical function. Additionally, research is required to understand
298 the cause of the reported difficulties in completing ADLs. While the ADL questions may not
299 detect decreased physical function, the responses should not be ignored or dismissed.

300 Future research that includes P_{sts} should examine several additional factors that impact
301 physical function. This research's nutritional status was limited to BMI, but it should include
302 body composition. Physical function is both a cause and a consequence of lean body mass,
303 making it an important consideration when evaluating physical function. Additionally, more
304 detailed dietary intake and biomarkers such as hemoglobin should be measured. Diet quality,
305 especially protein and micronutrient intake, impacts physical function (31–33). Physical
306 activity, including type, intensity, and duration, should be included in future research. Medical
307 history was limited to self-reported long or short-term illness or injury in this research, but an in-
308 depth medical history would be valuable to future research, including the type, timing, treatment,
309 and health-seeking behaviors around illnesses and injuries. Finally, research regarding physical
310 function needs more frequent data points over longer periods. There are seasonal variations in
311 many factors that influence physical function, including diet (57), physical activity (46), and
312 exposure to diseases (58). More data points will help determine how quickly physical function
313 changes in women in rural Ethiopia in response to different factors and if there are long-term
314 changes.

315 The STS was piloted among the enumerators before the start of the study. There was not
316 a consistent chair height among the households of the study participants, eliminating the
317 possibility of using a chair for the seated position. Like other cultures, people in rural Ethiopia
318 frequently sit in a deep squatted position while resting or working near the ground (e.g., farming
319 activities such as planting and weeding or domestic work such as cooking or laundry). As such,
320 the study team determined that it would be an appropriate option for the seated position of the
321 STS, especially since the P_{sts} variable took leg length into account. The enumerators determined
322 that seven was the ideal number of repetitions. It was challenging enough to reflect variations in
323 physical function, but participants could complete it.

324 **Conclusions**

325 Measuring physical function and identifying those experiencing difficulties completing
326 their ADLS is essential in all vulnerable populations, including women in rural Ethiopia. ADLS
327 in rural highland Ethiopia are physically demanding. Taking care of oneself, family, and
328 household, and many income-generating activities require significant manual labor and traveling
329 significant distances on foot. Therefore, decreases in physical function could affect the well-
330 being of one's family by impacting household income, diet quality, and health through water,
331 sanitation, and hygiene.

332 A feasible and reliable measure of physical function is vital in addressing poverty and
333 malnutrition. It would allow healthcare providers to identify those requiring additional medical
334 care and researchers to identify factors that influence physical function, including nutritional,
335 behavioral, and environmental factors. This research has demonstrated that the P_{sts} using a timed
336 seven-repetition stand to seated squat is a reliable measure of physical function in this setting.

337 **Acknowledgements**

338 We would like to thank our field research staff in Tigray, Ethiopia, especially Equbay
339 Gebrehiwet, for all their work during data collection. Additionally, we are thankful for the
340 support provided by the Ethiopian Strategic Support of the International Food Policy Research
341 Institute. We would like to thank Drs Kathleen M. Rasmussen and James Booth for their
342 support.

343 **References**

- 344 1. Spector WD, Katz S, Murphy JB, Fulton JP. The hierarchical relationship between activities of
345 daily living and instrumental activities of daily living. *Journal of Chronic Diseases*. 1987 Jan
346 1;40(6):481–9.
- 347 2. Beaton DE, Schemitsch E. Measures of health-related quality of life and physical function.
348 *Clinical Orthopaedics and Related Research*®. 2003 Aug;413:90–105.
- 349 3. Tomey KM, Sowers MR. Assessment of Physical Functioning: A Conceptual Model
350 Encompassing Environmental Factors and Individual Compensation Strategies. *Phys Ther*.
351 2009 Jul;89(7):705–14.
- 352 4. O’Neill D, Forman DE. The importance of physical function as a clinical outcome:
353 Assessment and enhancement. *Clin Cardiol*. 2020;43(2):108–17.
- 354 5. Seeman TE, Charpentier PA, Berkman LF, Tinetti ME, Guralnik JM, Albert M, et al. Predicting
355 Changes in Physical Performance in a High-Functioning Elderly Cohort: MacArthur Studies of
356 Successful Aging. *J Gerontol*. 1994 May 1;49(3):M97–108.
- 357 6. Wolinsky FD, Miller DK, Andresen EM, Malmstrom TK, Miller JP, Miller TR. Effect of
358 subclinical status in functional limitation and disability on adverse health outcomes 3 years
359 later. *J Gerontol A Biol Sci Med Sci*. 2007 Jan;62(1):101–6.
- 360 7. Viccaro LJ, Perera S, Studenski SA. Is Timed Up and Go Better Than Gait Speed in Predicting
361 Health, Function, and Falls in Older Adults? *J Am Geriatr Soc*. 2011 May;59(5):887–92.
- 362 8. Shukla B, Bassement J, Vijay V, Yadav S, Hewson D. Instrumented Analysis of the Sit-to-
363 Stand Movement for Geriatric Screening: A Systematic Review. *Bioengineering*. 2020
364 Dec;7(4):139.
- 365 9. Gleeson M, Sherrington C, Keay L. Exercise and physical training improve physical function
366 in older adults with visual impairments but their effect on falls is unclear: a systematic
367 review. *J Physiother*. 2014 Sep;60(3):130–5.

- 368 10. Fried TR, Bradley EH, Williams CS, Tinetti ME. Functional disability and health care
369 expenditures for older persons. *Arch Intern Med*. 2001 Nov 26;161(21):2602–7.
- 370 11. Morey MC, Sloane R, Pieper CF, Peterson MJ, Pearson MP, Ekelund CC, et al. Effect of
371 Physical Activity Guidelines on Physical Function in Older Adults. *Journal of the American*
372 *Geriatrics Society*. 2008;56(10):1873–8.
- 373 12. Beswick AD, Rees K, Dieppe P, Ayis S, Gooberman-Hill R, Horwood J, et al. Complex
374 interventions to improve physical function and maintain independent living in elderly
375 people: a systematic review and meta-analysis. *The Lancet*. 2008 Mar 1;371(9614):725–35.
- 376 13. Calvo-Ayala E, Khan BA, Farber MO, Ely EW, Boustani MA. Interventions to Improve the
377 Physical Function of ICU Survivors: A Systematic Review. *Chest*. 2013 Nov 1;144(5):1469–80.
- 378 14. Imamura K, Kamide N, Sakamoto M, Sato H, Shiba Y, Matsunaga A. Association Between
379 Social Network and Physical Function in Community-Dwelling Older Adults in Japan. *Phys*
380 *Ther Res*. 2020 Jul 22;23(2):153–9.
- 381 15. Rogers SN, Lowe D, Yueh B, Weymuller Jr EA. The Physical Function and Social-Emotional
382 Function Subscales of the University of Washington Quality of Life Questionnaire. *Arch*
383 *Otolaryngol Head Neck Surg*. 2010 Apr 19;136(4):352–7.
- 384 16. Noppert GA, Brown CS, Chanti-Ketterl M, Hall KS, Newby LK, Cohen HJ, et al. The Impact of
385 Multiple Dimensions of Socioeconomic Status on Physical Functioning Across the Life
386 Course. *Gerontol Geriatr Med*. 2018 Jan 1;4:2333721418794021.
- 387 17. Nilsen C, Agahi N, Kåreholt I. Work Stressors in Late Midlife and Physical Functioning in Old
388 Age. *J Aging Health*. 2017 Aug 1;29(5):893–911.
- 389 18. Palumbo A, De Roos A, Cannuscio C, Robinson L, Mossey J, Weitlauf J, et al. Work
390 Characteristics Associated with Physical Functioning in Women. *IJERPH*. 2017 Apr
391 15;14(4):424.
- 392 19. Paddick SM, Gray WK, Ogunjimi L, Iwezuala B, Olakehinde O, Kisoli A, et al. Validation of the
393 Identification and Intervention for Dementia in Elderly Africans (IDEA) cognitive screen in
394 Nigeria and Tanzania. *BMC Geriatr*. 2015 Dec;15(1):53.
- 395 20. Bucks RS, Ashworth DL, Wilcock GK, Siegfried K. Assessment of Activities of Daily Living in
396 Dementia: Development of the Bristol Activities of Daily Living Scale. *Age and Ageing*. 1996
397 Mar 1;25(2):113–20.
- 398 21. Fitzpatrick AL, Buchanan CK, Nahin RL, DeKosky ST, Atkinson HH, Carlson MC, et al.
399 Associations of Gait Speed and Other Measures of Physical Function with Cognition in a
400 Healthy Cohort of Elderly Persons. *The Journals of Gerontology: Series A*. 2007 Nov
401 1;62(11):1244–51.

- 402 22. Wang L, Larson EB, Bowen JD, van Belle G. Performance-Based Physical Function and Future
403 Dementia in Older People. *Archives of Internal Medicine*. 2006 May 22;166(10):1115–20.
- 404 23. Busch T de A, Duarte YA, Pires Nunes D, Lebrão ML, Satya Naslavsky M, dos Santos
405 Rodrigues A, et al. Factors associated with lower gait speed among the elderly living in a
406 developing country: a cross-sectional population-based study. *BMC Geriatr*. 2015 Apr
407 1;15:35.
- 408 24. Jones SE, Kon SSC, Canavan JL, Patel MS, Clark AL, Nolan CM, et al. The five-repetition sit-to-
409 stand test as a functional outcome measure in COPD. *Thorax*. 2013 Nov 1;68(11):1015–20.
- 410 25. White DK, Master H. Patient Reported Measures of Physical Function in Knee Osteoarthritis.
411 *Rheum Dis Clin North Am*. 2016 May;42(2):239–52.
- 412 26. Urits I, Smoots D, Franscioni H, Patel A, Fackler N, Wiley S, et al. Injection Techniques for
413 Common Chronic Pain Conditions of the Foot: A Comprehensive Review. *Pain Ther*. 2020
414 Jun;9(1):145–60.
- 415 27. Bernard C, Dabis F, de Rekeneire N. Physical function, grip strength and frailty in people
416 living with HIV in sub-Saharan Africa: systematic review. *Trop Med Int Health*. 2017
417 May;22(5):516–25.
- 418 28. Akosile CO, Mgbeojedo UG, Maruf FA, Okoye EC, Umeonwuka IC, Ogunniyi A. Depression,
419 functional disability and quality of life among Nigerian older adults: Prevalences and
420 relationships. *Archives of Gerontology and Geriatrics*. 2018 Jan;74:39–43.
- 421 29. Pieterse S, Manandhar M, Ismail S. The association between nutritional status and handgrip
422 strength in older Rwandan refugees. *Eur J Clin Nutr*. 2002 Oct;56(10):933–9.
- 423 30. Central Statistical Agency (CSA) Ethiopia and ICF. Ethiopia Demographic and Health Survey.
424 Addis Ababa, Ethiopia and Rockville, Maryland, USA: CSA and ICF; 2016.
- 425 31. McLoughlin DM, Spargo E, Wassif WS, Newham DJ, Peters TJ, Lantos PL, et al. Structural and
426 functional changes in skeletal muscle in anorexia nervosa. *Acta Neuropathol*. 1998 Jun
427 15;95(6):632–540.
- 428 32. Cimolin V, Galli M, Vismara L, Vimercati SL, Precilios H, Cattani L, et al. Gait Analysis in
429 Anorexia and Bulimia Nervosa. *Journal of Applied Biomaterials & Functional Materials*. 2013
430 May;11(2):122–8.
- 431 33. Asamane EA, Greig CA, Thompson JL. The association between nutrient intake, nutritional
432 status and physical function of community-dwelling ethnically diverse older adults. *BMC*
433 *Nutrition*. 2020 Aug 25;6(1):36.

- 434 34. Cesari M, Penninx BWJH, Pahor M, Lauretani F, Corsi AM, Williams GR, et al. Inflammatory
435 Markers and Physical Performance in Older Persons: The InCHIANTI Study. *The Journals of*
436 *Gerontology: Series A*. 2004 Mar 1;59(3):M242–8.
- 437 35. Brinkley TE, Leng X, Miller ME, Kitzman DW, Pahor M, Berry MJ, et al. Chronic Inflammation
438 Is Associated with Low Physical Function in Older Adults Across Multiple Comorbidities. *J*
439 *Gerontol A Biol Sci Med Sci*. 2009 Apr;64A(4):455–61.
- 440 36. Aiello AE, Haan MN, Pierce CM, Simanek AM, Liang J. Persistent Infection, Inflammation,
441 and Functional Impairment in Older Latinos. *The Journals of Gerontology: Series A*. 2008 Jun
442 1;63(6):610–8.
- 443 37. Wiener JM, Hanley RJ, Clark R, Van Nostrand JF. Measuring the Activities of Daily Living:
444 Comparisons Across National Surveys. *Journal of Gerontology*. 1990 Nov 1;45(6):S229–37.
- 445 38. Studenski S, Perera S, Wallace D, Chandler JM, Duncan PW, Rooney E, et al. Physical
446 Performance Measures in the Clinical Setting. *Journal of the American Geriatrics Society*.
447 2003 Mar;51(3):314–22.
- 448 39. Cesari M, Scarlata S. What the 4-metre gait speed measures and why it cannot replace
449 functional capacity tests. *European Respiratory Journal*. 2014 Jun 1;43(6):1819–20.
- 450 40. Lindemann U, Claus H, Stuber M, Augat P, Mucic R, Nikolaus T, et al. Measuring power
451 during the sit-to-stand transfer. *Eur J Appl Physiol*. 2003 Jun;89(5):466–70.
- 452 41. Smith WN, Rossi GD, Adams JB, Abderlahman K, Asfour SA, Roos BA, et al. Simple
453 equations to predict concentric lower-body muscle power in older adults using the 30-
454 second chair-rise test: a pilot study. *Clinical Interventions in Aging*. :8.
- 455 42. Guralnik JM, Seeman TE, Tinetti ME, Nevitt MC, Berkman LF. Validation and use of
456 performance measures of functioning in a non-disabled older population: MacArthur
457 studies of successful aging. *Aging Clin Exp Res*. 1994 Dec;6(6):410–9.
- 458 43. Curb JD, Ceria-Ulep CD, Rodriguez BL, Grove J, Guralnik J, Willcox BJ, et al. Performance-
459 Based Measures of Physical Function for High-Function Populations. *Journal of the*
460 *American Geriatrics Society*. 2006;54(5):737–42.
- 461 44. Katz S. Assessing Self-maintenance: Activities of Daily Living, Mobility, and Instrumental
462 Activities of Daily Living. *Journal of the American Geriatrics Society*. 1983;31(12):721–7.
- 463 45. Algarni AD, Alrabai HM, Al-Ahaideb A, Kachanathu SJ, AlShammari SA. Arabic translation,
464 cultural adaptation, and validation study of Knee Outcome Survey: Activities of Daily Living
465 Scale (KOS-ADLS). *Rheumatol Int*. 2017 Sep;37(9):1585–9.
- 466 46. Berhane G, Golan J, Hirvonen K, Hoddinott JF, Kim, Sunny, Seyoum Tafesse, Alemayehu, et
467 al. Evaluation of the nutrition-sensitive features of the fourth phase of Ethiopia’s Productive

- 468 Safety Net Programme [Internet]. 0 ed. Washington, DC: International Food Policy Research
469 Institute; 2020. Available from:
470 <https://ebrary.ifpri.org/digital/collection/p15738coll2/id/133685>
- 471 47. World Bank. The State of Social Safety Nets 2018. Washington, DC: World Bank; 2018.
472 Report No.: doi:10.1596/978-1-4648-1254-5.
- 473 48. Golan J, Hoddinott J. Validation of the 24-hour perceived exertion recall survey (PERS) in
474 women in rural Tigray, Ethiopia. *Current Developments in Nutrition*. 2023 Mar;100064.
- 475 49. Edwards MM. The Reliability and Validity of Self-Report Activities of Daily Living Scales.
476 *Canadian Journal of Occupational Therapy*. 1990;57(5):273–8.
- 477 50. Vitória Puzzi, Joice Mara Oliveira, Thainá Bessa Alves, Natielly Soares Correia, Diery
478 Fernandes Rugila, Thamires Spositon, et al. Validity and reproducibility of instruments to
479 assess activities of daily living in adults with asthma: preliminary results. *Eur Respir J*. 2021
480 Sep 5;58(suppl 65):PA761.
- 481 51. Pan AW, Wu CY, Chung L, Chen TJ. Reliability and validity of the self-reported Activities of
482 Daily Living Scale for people with mental illness. *Hong Kong J Occup Ther*. 2018
483 Dec;31(2):115–24.
- 484 52. Takai Y, Ohta M, Akagi R, Kanehisa H, Kawakami Y, Fukunaga T. Sit-to-stand Test to Evaluate
485 Knee Extensor Muscle Size and Strength in the Elderly: A Novel Approach. *J Physiol*
486 *Anthropol*. 2009;28(3):123–8.
- 487 53. Gurland BJ, Maurer MS. Life and Works of Sidney Katz, MD: A Life Marked by Fundamental
488 Discovery. *Journal of the American Medical Directors Association*. 2012 Nov;13(9):764–5.
- 489 54. Bull FC, Maslin TS, Armstrong T. Global Physical Activity Questionnaire (GPAQ): Nine
490 Country Reliability and Validity Study. *J Phys Act Health*. 2009 Nov;6(6):790–804.
- 491 55. Watson PF, Petrie A. Method agreement analysis: A review of correct methodology.
492 *Theriogenology*. 2010 Jun;73(9):1167–79.
- 493 56. Hoddinott J, Behrman JR, Maluccio JA, Melgar P, Quisumbing AR, Ramirez-Zea M, et al.
494 Adult consequences of growth failure in early childhood. *Am J Clin Nutr*. 2013
495 Nov;98(5):1170–8.
- 496 57. Baye K, Hirvonen K. Seasonality: a missing link in preventing undernutrition. *Lancet Child*
497 *Adolesc Health*. 2020 Jan 1;4(1):e3.
- 498 58. Altizer S, Dobson A, Hosseini P, Hudson P, Pascual M, Rohani P. Seasonality and the
499 dynamics of infectious diseases. *Ecology Letters*. 2006;9(4):467–84.

500