

RESEARCH ARTICLE

Assessing independence in mobility activities in trauma care: Validity and reliability of the Activity Independence Measure-Trauma (AIM-T) in humanitarian settings

Bérangère Gohy^{1,2*}, Christina H. Opava¹, Johan von Schreeb³, Rafael Van den Bergh⁴, Aude Brus⁵, Nicole Fouda Mbarga⁶, Jean Patrick Ouamba⁶, Jean-Marie Mafuko⁷, Irene Mulombwe Musambi⁸, Delphine Rougeon⁸, Evelyne Côté Grenier⁹, Lívia Gaspar Fernandes⁹, Julie Van Hulse¹⁰, Eric Weerts², The AIM-T Study Group^{4,6,7,8,9,10†}, Nina Brodin^{1,11}

1 Department of Neurobiology, Division of Physiotherapy, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden, **2** Humanity & Inclusion, Rehabilitation Technical Direction, Brussels, Belgium, **3** Department of Global Public Health, Karolinska Institutet, Stockholm Sweden, **4** Médecins Sans Frontières, Operational Center Brussels, Brussels, Belgium, **5** Humanity & Inclusion, Innovation, Impact & Information Division, Brussels, Belgium, **6** Médecins Sans Frontières, Operational Center Geneva, YaoundeYaounde, Cameroon, **7** Médecins Sans Frontières, Operational Center Brussels, Bujumbura, Burundi, **8** Médecins Sans Frontières, Operational Center Paris, Bangui, Central African Republic, Baghdad, Iraq, **9** Médecins Sans Frontières, Operational Center Paris, Baghdad, Iraq, **10** Médecins Sans Frontières, Operational Center Paris, France, **11** Department of Orthopaedics, Danderyd Hospital Corp., Division of Physiotherapy, Danderyd, Sweden

[†] Membership of the AIM-T study group is provided in the Acknowledgments

* berangere.gohy@ki.se



OPEN ACCESS

Citation: Gohy B, Opava CH, von Schreeb J, Van den Bergh R, Brus A, Fouda Mbarga N, et al. (2023) Assessing independence in mobility activities in trauma care: Validity and reliability of the Activity Independence Measure-Trauma (AIM-T) in humanitarian settings. *PLOS Glob Public Health* 3(9): e0001723. <https://doi.org/10.1371/journal.pgph.0001723>

Editor: Vinay Nair Kampalath, University of Pennsylvania Perelman School of Medicine, UNITED STATES

Received: February 7, 2023

Accepted: August 17, 2023

Published: September 11, 2023

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pgph.0001723>

Copyright: © 2023 Gohy et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

The importance of measuring outcomes after injury beyond mortality and morbidity is increasingly recognized, though underreported in humanitarian settings. To address shortcomings of existing outcome measures in humanitarian settings, the Activity Independence Measure-Trauma (AIM-T) was developed, and is structured in three subscales (i.e., core, lower limb, and upper limb). This study aimed to assess the AIM-T construct validity (structural validity and hypothesis testing) and reliability (internal consistency, inter-rater reliability and measurement error) in four humanitarian settings (Burundi, Iraq, Cameroon and Central African Republic). Patients with acute injury (n = 195) were assessed using the AIM-T, the Barthel Index (BI), and two pain scores. Structural validity was assessed through confirmatory factor analysis. Hypotheses were tested regarding correlations with BI and pain scores using Pearson correlation coefficient (PCC) and differences in AIM-T scores between patients' subgroups, using standardized effect size Cohen's d (d). Internal consistency was assessed with Cronbach's alpha (α). AIM-T was reassessed by a second rater in 77 participants to test inter-rater reliability using intraclass correlation coefficient (ICC). The results showed that the AIM-T structure in three subscales had an acceptable fit. The AIM-T showed an inverse weak to moderate correlation with both pain scores (PCC<0.7, p≤0.05), positive strong correlation with BI (PCC≥0.7, p≤0.05), and differed between all subgroups (d≥0.5, p≤0.05). The inter-rater reliability in the (sub)scales was good to excellent (ICC

Data Availability Statement: The authors confirm that, for approved reasons, some access restrictions apply to the data underlying the findings. Due to the sensitive nature of trauma care (including violent trauma) data, full datasets are not made available by default. Data are available through the MSF Data Sharing Agreement for researchers who meet the criteria for access to confidential data; requests should be addressed to the Data Sharing Agreement coordinator, Annick Antierens (Annick.Antierens@brussels.msf.org).

Funding: Elrha's Research for Health in Humanitarian Crises (R2HC) programme (Grant No.32398) funded the research coordination (BG) and research activity for this study, funded by the UK Foreign, Commonwealth and Development Office (FCDO), Wellcome Trust, and the UK National Institute for Health Research (NIHR) (<https://www.elrha.org/programme/research-for-health-in-humanitarian-crises/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

0.86–0.91) and the three subscales' internal consistency was adequate ($\alpha \geq 0.7$). In conclusion, this study supports the AIM-T validity in measuring independence in mobility activities and its reliability in humanitarian settings, as well as it informs on its interpretability. Thus, the AIM-T could be a valuable measure to assess outcomes after injury in humanitarian settings.

Introduction

Injury accounts for 4.3 million deaths yearly and represented 9.8% of the global burden of disease in 2019 [1]. While the years of life lost (YLLs) rate due to injury has decreased by 23.6% between 2010 and 2019, years lived with disability (YLDs) have increased by 1.1% over the same time period [2]. In low- and middle-income countries (LMICs), as well as in humanitarian settings, the burden of injury is high, though its extent is underestimated and underreported [3–6]. Humanitarian settings are defined here as “situations in which there is a widespread threat to life, physical safety, health or basic subsistence that is beyond the coping capacity of individuals and communities in which they reside”, due to chronic or sudden-onset crises, caused by natural or technological disasters, famine, epidemics or armed conflict [7]. In such settings, persons living with disability, either resulting from injury or from other health conditions, may face additional barriers in accessing basic needs (e.g., food or shelter), eventually suffering greater social and economic consequences [8–10]. The implementation of organized trauma care has significantly decreased mortality rates after injury [11]. Trauma care should however also strive to prevent avoidable disability in humanitarian settings [3].

Similarly, indicators used for trauma care monitoring have mainly focused on hospital processes and morbidity and mortality outcomes rather than on disability [5,12–15]. However, functioning has increasingly been recommended as a quality indicator of trauma care [14,16–19]. Full functioning and complete disability are extremes of the same continuum, capturing the “dynamic interaction between a person's health condition, environmental factors and personal factors” [20,21]. Although functioning and disability need to be comprehensively understood, targeted assessments of specific domains are recommended for patient-centered care [22,23]. Mobility and self-care activities are frequently limited after injury, and regaining independence in these activities is among the patients' priorities [24]. Their measurement is therefore an essential part of daily clinical practice, while also being useful for trauma care monitoring. In particular, the use of measures of independence in activities based on observation may increase opportunities to perform such activities within routine care, thereby encouraging early mobilization, which is a common challenge in humanitarian settings [19,25,26].

Several measures have been used to assess independence in activities after injury, including Barthel Index (BI), Functional Independence Measure (FIM), Short Musculoskeletal Function Assessment Questionnaire (SMFA), Short Form health survey–Physical function (SF36-PF), Activity Measure for Post-Acute Care (AM-PAC) and Patient-Reported Outcome Measurement Information System–Physical function (PROMIS-PF) [27–32]. Most of these have been developed and assessed in non-humanitarian settings. Some measures are self-reported, which may be impractical in settings with low literacy rate [33]. Also, lengthy questionnaires, accredited administration or license fees, lack of available versions in local languages, or lack of culturally adapted content are barriers that limit application of existing measures across humanitarian settings and hinder comparability of results [33–37]. Furthermore, in heterogeneous populations, such as patients after injury, the use of multiple specific measures hinders comparison between subgroups and increases clinician workload [38]. The Activity

Independence Measure–Trauma (AIM-T) has the potential of addressing a number of shortcomings of existing measures, being a generic and observed measure designed for and used in humanitarian settings. A preliminary version of this measure was described in 2016 as part of retrospective study in Afghanistan and was subsequently revised to address time constraints, cultural relevance, and appropriateness, based on content validity assessment [39,40]. The quality of the information produced by a measure depends on its measurement properties (i.e., validity and reliability) in the context where it is used [41]. To further support the AIM-T validity as a measure of independence and its reliable use across health care professionals and health structures, our study aimed to assess the AIM-T construct validity (structural validity and hypothesis testing) and reliability (internal consistency, inter-rater reliability, and measurement error) among patients admitted for trauma care in four humanitarian settings. Testing the AIM-T measurement properties in its context of use allows informed choices by potential users, eventually fostering the reporting of functioning after injury in such settings.

Materials and methods

This cross-sectional study was conducted between July 2019 and November 2019 and tested the revised version of the AIM-T, according to the Consensus-based Standards for the selection of health Measurement Instruments framework (COSMIN) [42].

Study setting

We collected data from centres supported or run by the medical non-governmental organisation Médecins Sans Frontières (MSF), which provides trauma care in humanitarian settings. The selection of the four centres was based on their geographic diversity and their capacity to collect data.

Data from the following centres were included in this study:

- The MSF trauma centres of Arche (Bujumbura, Burundi) and Sica (Bangui, Central African Republic), which are set up following MSF trauma centre standards [43,44].
- The Regional Hospital of Maroua (Cameroon), where MSF supported the surgical department for urgent surgical cases.
- The MSF Baghdad Medical Rehabilitation Centre (Iraq), providing post-operative care.

In- and outpatient physiotherapy was provided in all four centres.

Study population

We aimed to include 50 consecutive patients in each of the four centres. Patients aged five years or above, within six months of their injury, and receiving in- or outpatient physiotherapy service were eligible. To ensure sufficiently large samples for each sub-group analysis, we balanced between centres receiving more patients with acute (i.e., within 30 days after injury) versus post-acute injury (i.e., between 31 days and 6 months after injury). Orthopaedic, visceral, and soft tissue injuries were grouped by location: lower limb and pelvis, upper limb, and trunk (i.e., spine, abdomen, and chest). Patients with isolated central neurologic injuries were excluded.

For inter-rater reliability assessment, a subsample of 20 patients was purposively selected in each centre for a second assessment. Selection aimed at diversity in terms of sex and age, as well as location, nature, severity, and acuteness of injury. Two raters were recruited from each centre. All raters were trained physiotherapists working in the study centres.

All participants consented to participate in the study. For patients under 18 years old or having cognitive difficulties, a legal representant gave written informed consent. Patients between 12 and 18 were also requested to provide their written informed assent to participate in the study. A witness was used to confirm verbal consent for patients with low literacy levels as well as for patients volunteering for the study but unable to give written consent.

Measurements

Routine and study-specific data were collected in the four study centres. Routine data included demographic information (age and sex), and clinical data on the injuries (date, location, nature, severity). The South African Triage Score (SATS) categorises the injury severity by colour, from minor injuries, labelled as ‘green’, to emergency to be seen immediately, labelled as ‘red’ [45–47]. The SATS was routinely scored and used as a proxy for injury severity in this study. Study-specific data comprised:

The *Activity Independence Measure–Trauma (AIM-T)* is composed of 12 activities grouped into three subscales (i.e., core, lower limb, upper limb), ranging from 0 to 10 for the core subscale and 0 to 25 for the lower limb and the upper limb subscales. The total score ranges from 0 to 60, based on the difficulties observed and level of human or material assistance required (higher score indicating a higher independence in activities) (S1 Fig).

The *Barthel Index (BI)* is a generic measurement of independence in activities, used for self-report or clinician rating [27]. It is composed of ten activities, with a total score ranging from 0 to 100, depending on the level of assistance required (higher score indicating a higher level of independence). The BI has been used in patients after injury, including in humanitarian settings [48–52]. The French and Arabic validated versions were used and were revised following a linguistic validation process instructed by BI license holders.

The *Visual Analogue Scale (VAS)* and *Faces Pain Scale-revised (FPS-R)* are self-reported pain measurements previously used with patients after injury [53,54]. To accommodate cultural and literacy aspects, patients were given the choice between the two scales [55–58]. Using either of these scales, pain was evaluated both at rest and during one activity chosen by each patient. The FPS-R scores were converted to equal VAS scores with “10” representing the worst pain and “0” no pain [57].

Procedure

An eight-hour rater training was provided remotely by the first author in French for the centres in Burundi, Cameroon and CAR, and in English translated to Arabic for the centre in Iraq. It included instructions on the use of the different measurements, and two two-hour training sessions on the AIM-T, using videos of each activity throughout the trainings to reduce potential information loss due to translation. A written guideline describing in detail the AIM-T activities and scoring system was also provided, together with feedback sessions on its use with pilot patients. Sessions were organised joining several centres when possible.

Study-specific data were collected by one rater in each centre, assessing, in sequential order, the included patients with the AIM-T, the BI, and either of the pain scales. The administration of the AIM-T and BI was timed and both were administered as observed measures. Self-reporting for BI was used only for patients refusing to perform any of its activities [59]. The mode of administration was documented for each of the BI activities.

To assess inter-rater reliability, the selected patients were re-assessed with the AIM-T by a second rater blinded to the first assessment results. A 30-minute rest was allowed between the two assessments. Test conditions were similar for both assessments in terms of equipment, instructions, and planned physiotherapy session.

The raters also documented any other relevant observation regarding the assessment conditions or patient health status (e.g., refusal to perform any activity or pain reported during the activity).

Data analyses

Data are described using frequencies, medians with interquartile ranges (IQR), or means and standard deviations (SD). All quantitative analyses were performed in SPSS version 27.

The construct validity of the AIM-T was assessed through *structural validity* and *hypothesis testing*. All hypotheses were formulated before data analysis. Confirmation of at least 75% of the hypotheses is considered sufficient to support construct validity [60].

Structural validity was investigated through confirmatory factor analysis (CFA). As suggested by the AIM-T content validity assessment, clinical reasoning, and literature, we hypothesized that a three-factor model, which represent the three AIM-T subscales (i.e., core, lower limb and upper limb activities), would better fit the items of the AIM-T than a single-factor model [39,61]. The three-factor and the single-factor models were both tested, using CFA in Jamovi 2.2.5. Model fits were assessed using three fit indices and their goodness of fit criteria: root mean square error of approximation (RMSEA) (<0.06), comparative fit index (CFI) (>0.95), and standardized root mean square residuals (SRMR) (≤ 0.08) [62]. At least one of these criteria should be met to support the structural validity [62]. The unidimensionality of the subscales was assessed using Principal Component Analysis (PCA). The (sub)scale was considered unidimensional when 1) the first factor explained $\geq 20\%$ of the variance and 2) the ratio of the variance explained by the first factor to the variance explained by the second factor is greater than four [63].

For *hypothesis testing*, we investigated *known-group validity* and *concurrent validity*. For *known-group validity*, hypotheses were formulated based on clinical experience and known associations of covariates (i.e., acuteness, location and number of injuries) with independence in activities after injury [64–68]. Patients' subgroups were expected to differ in AIM-T scores by a Cohen's *d* standardised effect sizes larger or equal to 0.5 [69]. For *concurrent validity*, hypotheses were tested using Pearson correlation coefficient (PCC) [70]. We hypothesized a weak to moderate negative correlation between the AIM-T and pain scores ($PCC < -0.7$), stronger when assessed during activity. Although pain and independence in activities are distinct constructs, they are commonly associated after injury, with pain interfering with daily activities [71,72]. We hypothesised a strong positive correlation ($PCC \geq 0.7$) between BI and the AIM-T because they intend to measure the same construct. The reliability of the AIM-T was assessed through *internal consistency*, *inter-rater reliability* and *measurement error*.

For *internal consistency*, a Cronbach alpha coefficient (α) was calculated for each (sub)scale having a confirmed unidimensionality in CFA. A coefficient equal to or larger than 0.70 was considered adequate [62]. If lower than 0.70, the effect of deleting separate items from the AIM-T to reach a higher α was investigated [73].

The *inter-rater reliability* was determined using the intraclass correlation coefficients (ICC) between pooled scores from the first and second raters. The ICC for absolute agreement was calculated for each (sub)scale, based on a two-way mixed effect [73]. Poor, moderate, good and excellent reliability were indicated by an ICC of less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.9, respectively [74]. An ICC greater than 0.7 indicating sufficient reliability based on COSMIN guidance [62,74]. Patients who refused to perform at least one activity were removed from the inter-rater reliability analysis.

Measurement error was calculated as the standard error of measurement (SEM), the smallest detectable change (SDC) and the limits of agreement (LoA) for the total score and subscales

scores. $SEM_{\text{agreement}}$ was derived from the ICC, as the root square of the error variance [73]. The SDC was calculated as $1.96 \times SD_{\text{difference}}$. For the LoA, a Bland and Altman plot was performed, which plots the mean difference between the first and the second raters against the pooled AIM-T means from the two raters [73,75]. Upper and lower LoA are defined as: $Mean_{\text{difference}} \pm SDC$. The proportion of the mean differences being within the limits of agreement was assessed, as well as the heteroscedasticity of the variables [73,76].

Ethics

The protocol of this study was approved by the MSF Ethics Review Board, Geneva, Switzerland (reference ID 1893) and the Swedish Ethical Review Authority, Stockholm, Sweden (Dnr 2022-02806-01). National ethics review committees competent for each participating centre also approved the study, i.e., Burundi National Ethics Committee for the Protection of Human Rights of Participants in Biomedical and Behavioural Research (15/04/2019), Scientific Committee for the Validation of Study Protocols and Research Results on Health in Central African Republic (28/UB/FACSS/CSCVPER/19), the National Ethics Committee for Research in Human Health in Cameroon (2019/08/1184/CE/CNERSH/SP) and ethics committees of the Baghdad Directorate of Health in Iraq (p. 1/5/10).

Results

Participants

A total of 195 patients were included from the four study centres: 50 in Baghdad, 44 in Bangui, 50 in Bujumbura, and 51 in Maroua. Most patients were young males (66% male, 62% aged between 18 and 45 years). Most patients presented with single injury (75%) and at least one lower limb injury (75%), while less patients presented with upper limb injury (30%), trunk injury (10%) or both upper and lower limb injury (10%). The median time since injury was 36 days (IQR 5–83). The characteristics of the included patients are presented in Table 1.

The mean AIM-T total score was 44.2 (SD +/- 10.5), while the mean BI score was 80.9 (SD +/- 18.6). It took on average 10.9 (SD +/- 10.2) and 10.5 minutes (SD +/- 9.8) to administer the AIM-T and the BI, respectively. The assessment of the BI was mostly self-reported (66%), with self-care activities (i.e., 'getting on and off toilet', 'bathing self', and 'dressing and undressing') being self-reported by at least 80% of the patients, rather than observed. Patients reported significantly more intense pain during activity (VAS mean 4.2, SD +/- 2.1 and FPS-R mean 5.7, SD +/- 2.5) than at rest (VAS mean 1.9, SD +/- 1.8 and FPS-R mean 2.7, SD +/- 2.7), $p < 0.001$.

Validity

For *structural validity*, the three-factor model met the COSMIN fit criteria and fit better to this study sample than the single-factor model (Table 2). The visual representation of the three-subscale model and the coefficient of each item is available in S2 Fig. The PCA supported the unidimensionality of the three subscales but not of the total scale. The first factor of the (sub) scales explained at least 20% of the variance, ranging from 38.9% for the total scale to 85.7% for the core subscale. The ratio of variance was above four for the three subscales (core 5.9, lower limb 7.5, and upper limb 7.6) but not for the total scale (1.1).

For *hypothesis testing*, in terms of *known group validity*, all the patients' subgroups differed significantly in AIM-T total or subscale scores with a magnitude equal or superior to the hypothesised differences ($d \geq 0.5$). The *concurrent validity* of the AIM-T was supported by its correlation with all the comparator measurements with the strength hypothesized ($PCC < 0.7$ for pain scales, and $PCC \geq 0.7$ for BI). AIM-T mean scores and standardised effect sizes per

Table 1. Characteristics of the 195 patients assessed to test validity and reliability of the Activity Independence Measure–Trauma (AIM-T), and of the 77 patients among them who were assessed a second time to test inter-rater reliability.

Characteristics	Total sample n (%)	Inter-rater reliability sample n (%)
Total number	195	77
Age, years		
5–17	32 (16.4)	13 (16.9)
18–45	121 (62.1)	42 (54.5)
>45	42 (21.5)	22 (28.6)
Sex		
Male	129 (66.2)	42 (54.5)
Female	66 (33.8)	35 (45.5)
SATS^a triage colour		
Green	13 (6.7)	10 (13.0)
Yellow	88 (45.1)	32 (41.5)
Orange	43 (22.1)	15 (19.5)
Red	1 (0.5)	0 (0.0)
Missing	50 (25.6)	20 (26.0)
Injury acuteness		
< 30 days	91 (46.7)	36 (46.8)
30 days–6 months	104 (53.3)	41 (53.2)
Injury type^b		
≥ 1 lower limb fracture	105 (53.5)	41 (53.2)
≥ 1 lower limb soft tissue injury	70 (35.9)	23 (29.9)
≥ 1 lower limb amputation	8 (4.1)	5 (6.5)
≥ 1 upper limb fracture	41 (21.0)	15 (19.5)
≥ 1 upper limb soft tissue injury	22 (11.8)	6 (7.8)
≥ 1 upper limb amputation	4 (2.1)	2 (2.6)
≥ 1 visceral injury	9 (4.6)	3 (3.9)
≥ 1 other injury	28 (14.3)	12 (15.6)
Number of Injuries		
Single	147 (75.4)	60 (77.9)
Multiple	48 (24.6)	17 (22.1)

^aSouth African Triage Score (SATS),

^bPatients could have more than one injury type.

<https://doi.org/10.1371/journal.pgph.0001723.t001>

Table 2. Goodness of fit indices for the AIM-T three-factor and single-factor models, based on structural equation modelling, comparing the COSMIN fit criteria to the observed fit indices to test structural validity.

	RMSEA ^a	CFI ^b	SRMR ^c	COSMIN criteria fulfilled (≥ one goodness of fit)
Goodness of fit criteria [62]	<0.06	>0.95	<0.08	NA
Three subscale Model	0.14	0.90	0.07	Yes
One factor Model	0.31	0.47	0.26	No

^aRMSEA = Root mean square error of approximation,

^bCFI = Comparative fit index (CFI),

^cStandardized root mean square residuals (SRMR).

<https://doi.org/10.1371/journal.pgph.0001723.t002>

Table 3. AIM-T mean scores (SD) of the total sample and the different subgroups, comparing the hypothesized and observed effect sizes to test known group validity.

Subgroups	n	Hypothesized Cohen's d	AIM-T score, mean (SD)	Observed Cohen's d	p value	Hypothesis confirmed
Total sample	195	NA	Total score: 44.2 (10.5)	NA	NA	NA
	195		Core score: 8.3 (2.3)			
	195		Lower limb score: 14.5 (7.7)			
	195		Upper Limb score: 21.4 (5.8)			
		≥0.5	Total score: 41.4 (11.7)	0.5	0.021	Yes
Acute injury	91	46.6 (8.6)				
Post-acute injury	104					
		≥0.5	Lower limb score: 11.9 (6.5)	1.3	0.047	Yes
≥ 1 lower limb injury	147	22.2 (5.8)				
No lower limb injury	48					
		≥0.5	Upper limb score: 16.0 (6.9)	1.3	<0.001	Yes
≥ 1 upper limb injury	59	23.8 (3.2)				
No upper limb injury	136					
		≥0.5	Total score: 46.0 (8.9)	0.7	<0.001	Yes
Single	147	38.7 (12.8)				
Multiple	48					

<https://doi.org/10.1371/journal.pgph.0001723.t003>

patients' subgroups are presented in Table 3, and the AIM-T correlation with comparator measurements in Table 4.

Reliability

The *internal consistency* was adequate for the three unidimensional subscales (α 0.82 for core, 0.92 for lower limb and 0.91 upper limb subscales respectively).

For *inter-rater reliability* testing, 6 (all from Bangui) of 83 re-assessed patients were removed from the analysis because they had refused to perform at least one activity during the second assessment. The remaining 77 patients were mostly young males with single limb injuries (Table 1). The mean AIM-T scores for the second raters was 41.9 (SD +/- 11.2), differing

Table 4. Correlation between the AIM-T and the Barthel Index and pain scores, compared with the hypothesized correlations to test concurrent validity.

Comparator measurement	n	Hypothesized correlation, r	Observed correlation, r	p value	95% Confidence interval	Hypothesis confirmed
Barthel Index	195	≥ 0.7	0.83	<0.001	0.78–0.87	Yes
Pain at rest						
VAS ^a	96	< -0.7	-0.26	0.010	-0.44 --0.06	Yes
FPS-R ^b	100	< -0.7	-0.47	<0.001	-0.60 --0.29	Yes
Pain during activity						
VAS ^a	89	< -0.7 and stronger than pain at rest	-0.32	0.002	-0.50 --0.12	Yes
FPS-R ^b	100		-0.52	<0.001	-0.65 --0.36	Yes

^aVAS = Visual analogue scale,

^bFPS-R = Faces pain scale–revised.

<https://doi.org/10.1371/journal.pgph.0001723.t004>

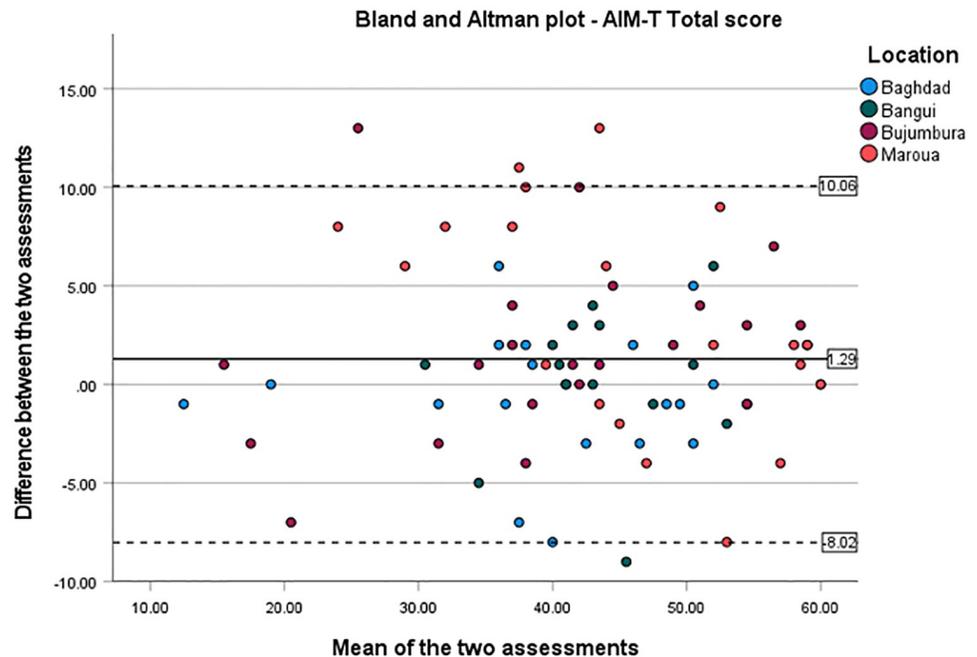


Fig 1. Bland and Altman plot for the AIM-T total score. Differences in AIM-T total scores between the first and the second raters are plotted against the pooled mean of the AIM-T scores for patients after injury in four humanitarian settings ($n = 77$). Mean difference between the two raters (-1.29) is represented by a solid line, while the limits of agreement (-8.02 to 10.07) by dashed lines.

<https://doi.org/10.1371/journal.pgph.0001723.g001>

significantly from the first raters on average by -1.3 ($SD \pm 4.6$) ($p = 0.008$, $CI 0.2$ to 2.3). The inter-rater reliability was excellent for the total scale ($ICC 0.91$, $CI 0.85$ to 0.94) and the lower limb subscale ($ICC 0.91$, $CI 0.86$ to 0.94), and good for the core ($ICC 0.86$, $CI 0.79$ to 0.91), and upper limb ($ICC 0.88$, $CI 0.81$ to 0.92) subscales.

Regarding *measurement error*, the $SEM_{agreement}$ was 3.36 , 0.96 , 2.41 , and 1.88 for the AIM-T total, core, lower limb, and upper limb (sub)scales respectively. The SDC was 9.31 , 2.66 , 6.68 , and 5.19 for the total, core, upper limb, and lower limb (sub)scales. The LoA ranged from -8.02 to 10.07 for the AIM-T total scale. As shown on the Bland and Altman plot, the measurement error between raters does not vary depending on the AIM-T scores, indicating no heteroscedasticity (Fig 1). The Bland and Altman plots for the subscales can be found in S3 Fig.

Discussion

This study supports the validity and reliability of the AIM-T in measuring independence in activities of patients after injury in four different humanitarian settings. First, different aspects of validity of the AIM-T are supported based on the consistence with pre-specified hypotheses concerning its structure, subgroup differences, and correlation with other validated measurements, exceeding the 75% COSMIN criteria [60]. Second, the AIM-T demonstrated good to excellent inter-rater reliability and adequate internal consistency, meeting COSMIN criteria to support its reliability [62].

AIM-T validity

The structural validity of the AIM-T composed of three subscales, core, lower limb and upper limb, is supported by our results. Indeed, mobility activities may involve the central, lower and

upper body in isolation or in combination, which may result in subdomains of mobility [61,77,78]. Subdomains have also been identified within existing unidimensional measures or item banks, reflecting either mobility versus self-care activities within the BI, the motor FIM or the AM-PAC, or lower versus upper limb activities with the PROMIS-PF, SMFA and motor FIM [79–85]. This study confirms that the AIM-T subscale scores provide more accurate information than the total score, thus supporting their use where feasible.

The AIM-T reflected the expected differences in independence in activities among patients after injury, thereby supporting the AIM-T known-group validity. Specifically, the clear difference between patients with upper versus lower limb injuries is consistent with differences documented by other measures, and further confirms the relevance of the AIM-T subscales [68,86]. The acuteness of injury was discriminated significantly by the AIM-T scores, which is coherent with differences identified by other measures as well as with the expected recovery of independence with time [87–89]. Additionally, the moderate observed difference ($d = 0.5$) might be related to the heterogeneity of patients as well to some patients having received continued reconstructive care, potentially prolonging their acute phase [90]. The ability of the AIM-T to discriminate acuteness longitudinally needs to be further investigated through responsiveness studies.

The strong correlation with the BI supports the concurrent validity of the AIM-T, and both measures require a similar administration time. The AIM-T's added value is that its activities can be performed similarly across patients, while being considered culturally appropriate to observe. Indeed, while most patients preferred to self-report the BI self-care activities, only few patients refused to perform any of the AIM-T activities, the reasons for which were pain or fear. Moreover, self-care activities may be performed differently across cultures and genders, leading to interpretation difficulties [34,36,91–93]. The AIM-T only includes mobility activities, having a more standardized performance. Additionally, the AIM-T instructions have been complemented with drawings to ease the understanding by patients and limit language barriers. Thus, by including only mobility activities that are culturally appropriate and standardized, comparability of results across settings should be facilitated. However, this may need further investigation since cross-cultural validity of the AIM-T was not assessed in the present study.

AIM-T reliability

The reliability of the AIM-T for patients after injury was supported for both the total scale and the subscales, through the testing of different measurement properties. This study shows that the items comprised within each AIM-T subscale consistently measure a defined construct, namely the independence in core, lower and upper limb activities [73]. The high values observed for the upper and lower limb subscales ($\alpha > 0.9$) are consistent with other measures assessing similar constructs [87,89,94–96]. The good to excellent inter-rater reliability reaches the COSMIN criteria ($ICC > 0.70$). This is comparable to that of other validated measurements, even though information in similar populations is lacking [79,97,98]. The AIM-T scores from second raters were systematically lower than those of first raters. This might suggest that patients have been too tired to perform the activities similarly or have experienced increased pain during the second assessments. The short time interval was chosen to ensure stability in health status of the patients, as done by others [98]. Moreover, the systematic difference between first and second raters is below the SEM, and therefore falls within the expected true scores.

Study strengths and limitations

This study has several strengths. Using the COSMIN framework ensured a rigorous methodology that enhances the use of the findings. Testing validity and reliability of a new measure is

part of its development process to ensure its adequacy [73]. To mitigate the risk for confirmation bias, data was collected by trained raters not involved in the study design or data analysis. An additional strength is the inclusion of a range of raters and a heterogeneous patient population, including children. Children are often excluded in trauma research despite bearing an important burden of injuries in humanitarian settings [99]. Also, the inclusion of pain assessment during activity and not just at rest allowed reflection of the correlation between pain and activity independence. This is in line with recommendations for assessing pain intensity in acute pain management, such as after surgery or trauma [100–102]. Testing the AIM-T against a widely used and validated measure such as the BI adds to current literature as it has been used in similar settings [48–52,103]. Besides, the BI was chosen as comparator measure due to its brevity, simplicity, and affordability. Lastly, the Cohen's *d* was used to assess the magnitude of the difference in AIM-T scores between patient subgroups, while others have used the SEM or not specified the magnitude [31,68,87]. Though COSMIN recommends to pre-specify and assess magnitude of differences, there is no guidance on assessment methods, complicating the comparability of results between studies. The use of the effect size allows to assess the magnitude of the differences between groups, as intended in construct validity testing [104].

This study has limitations, some connected to the challenges of conducting research in humanitarian settings. One limitation included the use of four different pairs of raters for assessing inter-rater reliability. This was done to reduce the burden of data collection and to ensure diversity of settings. To mitigate variability between pairs, all raters were trained by the first author using the same training package [97]. Another possible limitation was giving each participant the option to choose one out of two different pain scales (i.e., VAS and FPS-R), based on acceptability of the scales by the different providers [58,105]. There are contradictory findings regarding the strength of the correlation between the pain scales, hindering their interchangeability [57,58,106]. However, since all four correlation coefficients with the AIM-T are coherent, we do not consider this to have a major impact on our findings. Lastly, regarding structural validity, the three-factor model showed a limited goodness of fit, reaching only one fit criterion out of three. However, similar values have been considered acceptable by others, and only one fit criterion is required by COSMIN framework to support structural validity [62,107]. Furthermore, the measure was only tested in MSF-run or -supported health facilities, which could limit external validity. However, this is partly mitigated by the inclusion of four different humanitarian settings, located in different geographical regions.

Clinical implications and future research

Overall, more high-quality rehabilitation research has been called for, as well as strengthened health information systems, using data on functioning and rehabilitation, including in humanitarian settings [108]. The primary objective when developing the AIM-T was to provide a measure which could be routinely used by any healthcare professionals, quantifying the burden of injury beyond mortality and morbidity, while also being useful to guide clinical practice. Though it has been developed, used and tested within MSF projects, the AIM-T is intended to be useful in any health facilities providing trauma care, and especially where resources are limited. However, health authorities' awareness on the importance of collecting such data within trauma care is required, consistent with the more global call for stronger health information systems in LMICs [19,109,110]. Additionally, to ensure quality of data, a training package has been compiled for healthcare professionals, enhancing the AIM-T reliability and implementation. The package includes a guideline with a detailed description and illustration of each AIM-T activity, as well as a poster displaying the 12 activities and a flow-chart explaining the scoring system (S4 Fig). An open access e-learning module in English will

soon be available to enhance access to the training materials, while French and Arabic translations are planned in the near future. Having the SEM and SDC values of AIM-T informs healthcare professionals about its interpretability, but needs to be complemented with other reference values, such as its responsiveness, including the minimum important change. The use of the three subscales is recommended to better capture the independence and multidimensionality among patients after injury. To report the burden of injury, having only one total score to reflect independence in activities is favored, complying with operational needs and constraints. The unidimensionality of the AIM-T total scale was however not supported by our data and should be further investigated in a larger sample, comparing model fit with different profiles of patients. Additionally, we plan to conduct a longitudinal study using the AIM-T as one of the measures to document recovery of functioning after injury in different humanitarian settings and to identify its association with early rehabilitation. Further, describing the experience of healthcare professionals and patients with the use of the AIM-T to encourage early rehabilitation in such settings would be valuable.

Conclusion

This study supports the validity and reliability of the AIM-T as a measure of independence in activities in patients after injury in humanitarian settings, complementing previous findings on its validity. Thus, the AIM-T could be a valuable measure in trauma care to assess outcomes after injury in humanitarian settings, fostering the reporting of functioning and the monitoring of quality of trauma care in such settings.

Supporting information

S1 Fig. Activity Independence Measure–Trauma (AIM-T), the measure.
(PDF)

S2 Fig. The three-factor model of the AIM-T with factor loadings.
(PDF)

S3 Fig. Bland and Altman plots for the AIM-T subscales.
(PDF)

S4 Fig. Activity Independence Measure–Trauma (AIM-T), poster and scoring system flow-chart.
(PDF)

Acknowledgments

We would like to thank the local teams in each study centre, as well as the coordination and headquarters teams at Médecins Sans Frontières and Humanity & Inclusion. Despite working in challenging environments, their commitment and support made this study happen. More particularly, we would like to thank the eight raters as well as Geraldine Duc, Vincent Lambert, Khaled Ahmedana, Mohammed Abed, Zoe Clift and Jacinth Yan for their valuable contributions to this research project. We would also like to thank our study participants who took their time to participate on the data collection. This study has been funded by Elrha's Research for Health in Humanitarian Crises (R2HC) Programme, which aims to improve health outcomes by strengthening the evidence base for public health interventions in humanitarian crises. R2HC is funded by the UK Foreign, Commonwealth and Development Office (FCDO), Wellcome, and the UK National Institute for Health Research (NIHR).

The following co-authors are part of the AIM-T study group: Gedeon Fabrice Touye, Eric Ndiramiye, Iza Ciglenecki, Annick Antierens, and Arielle Calmejane.

Author Contributions

Conceptualization: Bérange Gohy, Christina H. Opava, Johan von Schreeb, Rafael Van den Bergh, Aude Brus, Julie Van Hulse, Eric Weerts, Nina Brodin.

Data curation: Bérange Gohy.

Formal analysis: Bérange Gohy, Nina Brodin.

Funding acquisition: Bérange Gohy, Christina H. Opava, Johan von Schreeb, Rafael Van den Bergh, Aude Brus, Julie Van Hulse, Eric Weerts, Nina Brodin.

Investigation: Bérange Gohy, Nicole Fouda Mbarga, Delphine Rougeon, Evelyne Côté Grenier, Lívia Gaspar Fernandes, Eric Weerts.

Methodology: Bérange Gohy, Christina H. Opava, Johan von Schreeb, Rafael Van den Bergh, Aude Brus, Julie Van Hulse, Eric Weerts, Nina Brodin.

Project administration: Bérange Gohy, Rafael Van den Bergh, Jean Patrick Ouamba, Jean-Marie Mafuko, Irene Mulombwe Musambi, Evelyne Côté Grenier, Julie Van Hulse, Eric Weerts, Nina Brodin.

Supervision: Bérange Gohy, Rafael Van den Bergh, Jean Patrick Ouamba, Jean-Marie Mafuko, Irene Mulombwe Musambi, Evelyne Côté Grenier, Julie Van Hulse, Eric Weerts, Nina Brodin.

Validation: Bérange Gohy, Christina H. Opava, Johan von Schreeb, Rafael Van den Bergh, Aude Brus, Julie Van Hulse, Eric Weerts, Nina Brodin.

Writing – original draft: Bérange Gohy, Christina H. Opava, Johan von Schreeb, Rafael Van den Bergh, Nina Brodin.

Writing – review & editing: Bérange Gohy, Christina H. Opava, Johan von Schreeb, Rafael Van den Bergh, Aude Brus, Nicole Fouda Mbarga, Jean Patrick Ouamba, Jean-Marie Mafuko, Irene Mulombwe Musambi, Delphine Rougeon, Evelyne Côté Grenier, Lívia Gaspar Fernandes, Julie Van Hulse, Eric Weerts, Nina Brodin.

References

1. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020; 396(10258):1204–22. [https://doi.org/10.1016/S0140-6736\(20\)30925-9](https://doi.org/10.1016/S0140-6736(20)30925-9) PMID: 33069326
2. Institute for Health Metrics and Evaluation. GBD 2019 Cause and Risk summaries [Internet]. 2020 [cited 2021 Apr 7]. http://www.healthdata.org/results/gbd_summaries/2019.
3. Sphere Association. The Sphere Handbook: Humanitarian Charter and Minimum Standards in Humanitarian Response [Internet]. Vol. 1, CHS Alliance, Sphere Association and Groupe URD. 2018. 402 p. www.practicalactionpublishing.org/sphere.
4. The Economist Intelligence Unit. At breaking point Understanding the impact of musculoskeletal injuries in low- and middle-income countries. 2018; <https://www.eiu.com/graphics/marketing/pdf/Injuries-in-LMICs.pdf>
5. Smith J, Roberts B, Knight A, Gosselin R, Blanchet K. A systematic literature review of the quality of evidence for injury and rehabilitation interventions in humanitarian crises. *Int J Public Health*. 2015; 60(7):865–72. <https://doi.org/10.1007/s00038-015-0723-6> PMID: 26298446
6. Haagsma JA, James SL, Castle CD, Dingels ZV., Fox JT, Hamilton EB, et al. Burden of injury along the development spectrum: Associations between the Socio-demographic Index and disability-

- adjusted life year estimates from the Global Burden of Disease Study 2017. *Inj Prev*. 2020; 26:i12–26. <https://doi.org/10.1136/injuryprev-2019-043296> PMID: 31915273
7. Elrha. Research Ethics Tool [Internet]. 2017. www.elrha.org/wp-content/uploads/2015/01/Evidence-Review-22.10.15.pdf.
 8. World Health Organization. Guidance Note on Disability and Emergency Risk Management for Health. World Health Organization. Geneva; 2013.
 9. Rathore FA, Gosney JE, Reinhardt JD, Haig AJ, Li J. Medical Rehabilitation After Natural Disasters: Why, When, and How? *Arch Phys Med Rehabil*. 2012; 93(10):1875–81.
 10. King J, Edwards N, Watling H, Hair S. Barriers to disability-inclusive disaster management in the Solomon Islands: Perspectives of people with disability. *Int J Disaster Risk Reduct*. 2019; 34:459–66.
 11. Moore L, Champion H, Tardif PA, Kuimi BL, O'Reilly G, Leppaniemi A, et al. Impact of Trauma System Structure on Injury Outcomes: A Systematic Review and Meta-Analysis. *World J Surg*. 2018; 42(5):1327–39. <https://doi.org/10.1007/s00268-017-4292-0> PMID: 29071424
 12. Gruen RL, Gabbe BJ, Stelfox HT, Cameron PA. Indicators of the quality of trauma care and the performance of trauma systems. *Br J Surg*. 2012; 99(SUPPL. 1):97–104. <https://doi.org/10.1002/bjs.7754> PMID: 22441862
 13. Juillard CJ, Mock C, Goosen J, Joshipura M, Civil I. Establishing the evidence base for trauma quality improvement: A collaborative WHO-IATSIC review. *World J Surg*. 2009; 33(5):1075–86. <https://doi.org/10.1007/s00268-009-9959-8> PMID: 19290573
 14. Rios-Diaz AJ, Herrera-Escobar JP, Lilley EJ, Appelson JR, Gabbe B, Brasel K, et al. Routine inclusion of long-term functional and patient-reported outcomes into trauma registries: The FORTE project. *J Trauma Acute Care Surg*. 2017; 83(1):97–104. <https://doi.org/10.1097/TA.0000000000001490> PMID: 28426563
 15. Segui-gomez M, Mackenzie EJ. Measuring the Public Health Impact of Injuries. *Epidemiol Rev*. 2003; 25(5):3–19. <https://doi.org/10.1093/epirev/mxg007> PMID: 12923986
 16. Stucki G, Bickenbach J. Functioning: The third health indicator in the health system and the key indicator for rehabilitation. *Eur J Phys Rehabil Med*. 2017; 53(1):134–8. <https://doi.org/10.23736/S1973-9087.17.04565-8> PMID: 28118696
 17. van der Vliet QMJ, Bhashyam AR, Hietbrink F, Houwert RM, Öner FC, Leenen LPH. Routine incorporation of longer-term patient-reported outcomes into a Dutch trauma registry. *Qual Life Res*. 2019; 28(10):2731–9. <https://doi.org/10.1007/s11136-019-02211-y> PMID: 31098797
 18. Dijkman BG, Sprague S, Schemitsch EH, Bhandari M. When is a fracture healed? Radiographic and clinical criteria revisited. *J Orthop Trauma*. 2010; 24 Suppl 1(3):S76–80. <https://doi.org/10.1097/BOT.0b013e3181ca3f97> PMID: 20182242
 19. Barth CA, Wladis A, Roy N, Blake C, Kolo SM, O'sullivan C. Ways to improve surgical outcomes in low-and middle-income countries. *Bull World Health Organ*. 2022; 100(11):726–32. <https://doi.org/10.2471/BLT.22.287950> PMID: 36324551
 20. World Health Organization. How to use the ICF—A Practical Manual for using the International Classification of Functioning, Disability and Health (ICF). Exposure draft for comment. Geneva; 2013.
 21. Cieza A, Sabariego C, Bickenbach J, Chatterji S. Rethinking Disability. *BMC Med*. 2018; 16(1):14. <https://doi.org/10.1186/s12916-017-1002-6> PMID: 29370847
 22. Üstün TB, Chatterji S, Bickenbach J, Kostanjsek N, Schneider M. The International Classification of Functioning, Disability and Health: A new tool for understanding disability and health. *Disabil Rehabil*. 2003; 25(11–12):565–71. <https://doi.org/10.1080/0963828031000137063> PMID: 12959329
 23. Stucki G, Grimby G. Foreword: Applying the ICF in medicine. *J Rehabil Med Suppl*. 2004;(44):5–6.
 24. Hoffman KP, Playford DE, Grill E, Soberg H, Brohi K. Minimum data set to measure rehabilitation needs and health outcome after major trauma: Application of an international framework. *Eur J Phys Rehabil Med*. 2016; 52(3):331–46. PMID: 26745362
 25. Kawamura K, Murayama K, Takamura J, Minegishi S. Effect of a weekly functional independence measure scale on the recovery of patient with acute stroke: A retrospective study. *Medicine (Baltimore)*. 2022; 101(11):e28974. <https://doi.org/10.1097/MD.00000000000028974> PMID: 35356904
 26. Gosling J, Golyk V, Mishra S, Skelton P. We must not neglect rehabilitation in Ukraine. *eClinicalMedicine*. 2022; 50:101537. <https://doi.org/10.1016/j.eclinm.2022.101537> PMID: 35812991
 27. Mahoney FI, Barthel DW. Functional Evaluation: The Barthel Index. *Md State Med J*. 1965; 14:56–61. PMID: 14258950
 28. Granger C V., Hamilton B, Keith R, Zielezny M, Sherwin FS. Advance in functional assessment for medical rehabilitation. *Top Geriatr Rehabil*. 1986; 1(3):59–74.

29. Swiontkowski MF, Engelberg R, Martin DP, Agel J. Short Musculoskeletal Function Assessment Questionnaire: Validity, reliability, and responsiveness. *J Bone Jt Surg.* 1999; 81(9):1245–60. <https://doi.org/10.2106/00004623-199909000-00006> PMID: 10505521
30. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. conceptual framework and item selection. *Med Care.* 1992; 30(6):473–83.
31. Jette DU, Stilphen M, Ranganathan VK, Passek SD, Frost FS, Jette AM. Validity of the AM-PAC “6-Clicks” Inpatient Daily Activity and Basic Mobility Short Forms. *Phys Ther.* 2014; 94(3):379–91. <https://doi.org/10.2522/ptj.20130199> PMID: 24231229
32. Ziedas AC, Abed V, Swantek AJ, Rahman TM, Cross A, Thomashow K, et al. Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Function Instruments Compare Favorably With Legacy Patient-Reported Outcome Measures in Upper- and Lower-Extremity Orthopaedic Patients: A Systematic Review of the Literature. *Arthrosc—J Arthrosc Relat Surg.* 2022; 38(2):609–31. <https://doi.org/10.1016/j.arthro.2021.05.031> PMID: 34052370
33. Reinhardt JD, Li J, Gosney J, Rathore FA, Haig AJ, Marx M, et al. Disability and health-related rehabilitation in international disaster relief. *Glob Health Action.* 2011; 4:7191. <https://doi.org/10.3402/gha.v4i0.7191> PMID: 21866223
34. Ali SM, Mulley GP. Is the Barthel Scale appropriate in non-industrialized countries? A view of rural Pakistan. *Disabil Rehabil.* 1998; 20(5):195–9. <https://doi.org/10.3109/09638289809166081> PMID: 9622265
35. World Confederation for Physical Therapy. WCPT report: The role of physical therapists in disaster management [Internet]. WCPT. London; 2016. <https://world.physio/sites/default/files/2020-06/Disaster-Management-Report-201603.pdf>.
36. Lundgren-Nilsson Å, Grimby G, Ring H, Tesio L, Lawton G, Slade A, et al. Cross-cultural validity of Functional Independence Measure items in stroke: A study using Rasch analysis. *J Rehabil Med.* 2005; 37(1):23–31. <https://doi.org/10.1080/16501970410032696> PMID: 15788329
37. van Zyl C, Mokkink LB, Derman W, Hanekom S, Heine M. Patient-Reported Outcome Measures in Key Sub-Saharan African Languages to Promote Diversity: A Scoping Review. *Value Heal Reg Issues.* 2023; 34:86–99. <https://doi.org/10.1016/j.vhri.2022.11.001> PMID: 36621214
38. Geraerds AJLM, Richardson A, Haagsma J, Derrett S, Polinder S. A systematic review of studies measuring health-related quality of life of general injury populations: Update 2010–2018. *Health Qual Life Outcomes.* 2020; 18(1):160. <https://doi.org/10.1186/s12955-020-01412-1> PMID: 32471430
39. Gohy B, Opava CH, Schreeb J Von, Van Den Bergh R, Brus A, Qaradaya AEH, et al. Monitoring independence in daily life activities after trauma in humanitarian settings: Item reduction and assessment of content validity of the Activity Independence Measure- Trauma (AIM-T). *PLOS Glob Public Heal.* 2022; 2(12):e0001334. <https://doi.org/10.1371/journal.pgph.0001334> PMID: 36962914
40. Gohy B, Ali E, Van den Bergh R, Schillberg E, Nasim M, Naimi MM, et al. Early physical and functional rehabilitation of trauma patients in the Médecins Sans Frontières trauma centre in Kunduz, Afghanistan: Luxury or necessity? *Int Health.* 2016; 8(6):381–9.
41. de Souza AC, Alexandre NMC, de Brito Guirardello E. Psychometric properties in instruments evaluation of reliability and validity. *Epidemiol Serv Saude.* 2017; 26(3):649–59.
42. Mokkink LB, Prinsen CA, Patrick D, Alonso J, Bouter LM, de Vet HC, et al. COSMIN Study Design checklist for Patient-reported outcome measurement instruments [Internet]. Amsterdam University Medical Centers. 2019. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf.
43. Valles P, Van den Bergh R, van den Boogaard W, Tayler-smith K, Gayraud O, Mammozai BA, et al. Emergency department care for trauma patients in settings of active conflict versus urban violence: all of the same calibre? *Int Health.* 2016; 8(6):390–7.
44. Hemat H, Shah S, Isaakidis P, Das M, Thu N, Kyaw T, et al. Before the bombing: High burden of traumatic injuries in Kunduz Trauma Center, Kunduz, Afghanistan. *PLoS One.* 2017; 12(3):1–9. <https://doi.org/10.1371/journal.pone.0165270> PMID: 28282377
45. Gottschalk SB, Wood D, Devries S, Wallis LA, Bruijns S. The cape triage score: a new triage system South Africa. *Emerg Med J.* 2006; 23:149–53.
46. Western Cape Government. The South African Triage Scale—Training manual 2012. Cape Town; 2012.
47. Dalwai MK, Tayler-Smith K, Trelles M, Jemmy JP, Maikéré J, Twomey M, et al. Implementation of a triage score system in an emergency room in Timergara, Pakistan [Short communication]. *Public Heal Action.* 2013; 3(1):43–5.
48. Zhang X, Reinhardt JD, Gosney JE, Li J. The NHV Rehabilitation Services Program Improves Long-Term Physical Functioning in Survivors of the 2008 Sichuan Earthquake: A Longitudinal Quasi Experiment. *PLoS One.* 2013; 8(1):1–10. <https://doi.org/10.1371/journal.pone.0053995> PMID: 23308293

49. Zhang J long, He H chen, Lin H dan, Luo Q lu, Xia L, Li S sha, et al. Motor function and activities of daily living capacity of patients with fractures sustained during the Wenchuan earthquake. *Chin Med J (Engl)*. 2011; 124(10):1504–7. PMID: [21740806](https://pubmed.ncbi.nlm.nih.gov/21740806/)
50. Castiglia SF, Galeoto G, Lauta A, Palumbo A, Tirinelli F, Viselli F, et al. The culturally adapted Italian version of the Barthel Index (IcaBI): Assessment of structural validity, inter-rater reliability and responsiveness to clinically relevant improvements in patients admitted to inpatient rehabilitation centers. *Funct Neurol*. 2017; 32(4):221–8. <https://doi.org/10.11138/fneur/2017.32.4.221> PMID: [29306359](https://pubmed.ncbi.nlm.nih.gov/29306359/)
51. Lorente S, Viladrich C, Vives J, Losilla JM. Tools to assess the measurement properties of quality of life instruments: A meta-review. *BMJ Open*. 2020; 10(8):e036038. <https://doi.org/10.1136/bmjopen-2019-036038> PMID: [32788186](https://pubmed.ncbi.nlm.nih.gov/32788186/)
52. Inui T, Watanabe Y, Kurata Y, Suzuki T, Matsui K, Ishii K, et al. Early postoperative Barthel index score and long-term walking ability in patients with trochanteric fractures walking independently before injury: a retrospective cohort study. *Arch Orthop Trauma Surg*. 2021; 141(8):1283–90. <https://doi.org/10.1007/s00402-020-03548-7> PMID: [32705380](https://pubmed.ncbi.nlm.nih.gov/32705380/)
53. Scott J, Huskisson EC. Graphic representation of pain. *Pain*. 1976; 2(2):175–1984. PMID: [1026900](https://pubmed.ncbi.nlm.nih.gov/1026900/)
54. Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The Faces Pain Scale—Revised: toward a common metric in pediatric pain measurement. *Pain*. 2001; 93(2):173–83. [https://doi.org/10.1016/S0304-3959\(01\)00314-1](https://doi.org/10.1016/S0304-3959(01)00314-1) PMID: [11427329](https://pubmed.ncbi.nlm.nih.gov/11427329/)
55. Huang KT, Owino C, Vreeman RC, Hagembe M, Njuguna F, Strother RM, et al. Assessment of the face validity of two pain scales in Kenya: A validation study using cognitive interviewing. *BMC Palliat Care*. 2012; 11(5).
56. Finianos J, Sanchez-Rodriguez E, Miro J. The Arabic Version of the Faces Pain Scale-Revised: Cultural Adaptation, Validity, and Reliability Properties When Used with Children and Adolescents. *Children*. 2021; 8(1184):1–8. <https://doi.org/10.3390/children8121184> PMID: [34943380](https://pubmed.ncbi.nlm.nih.gov/34943380/)
57. Le May S, Ballard A, Khadra C, Gouin S, Plint AC, Villeneuve E, et al. Comparison of the psychometric properties of 3 pain scales used in the pediatric emergency department. *Pain*. 2018; 159(8):1508–17.
58. Fadaizadeh L, Emami H, Samii K. Comparison of visual analogue scale and faces rating scale in measuring acute postoperative pain. *Arch Iran Med*. 2009; 12(1):73–5. PMID: [19111034](https://pubmed.ncbi.nlm.nih.gov/19111034/)
59. Roy CW, Togneri EH, Hay E, Pentland B. An inter-rater reliability study of the barthel index. *Int J Rehabil Res*. 1988; 11(1):67–70.
60. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007; 60(1):34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012> PMID: [17161752](https://pubmed.ncbi.nlm.nih.gov/17161752/)
61. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol*. 2008; 61(1):17–33. <https://doi.org/10.1016/j.jclinepi.2006.06.025> PMID: [18083459](https://pubmed.ncbi.nlm.nih.gov/18083459/)
62. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, De Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018; 27(5):1147–57.
63. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, et al. Psychometric Evaluation and Calibration of Health-Related Quality of Life Item Banks. *Med Care*. 2007; 45(5):S22–31.
64. De Munter L, Polinder S, Havermans RJM, Steyerberg EW, De Jongh MAC. Prognostic factors for recovery of health status after injury: A prospective multicentre cohort study. *BMJ Open*. 2021; 11(1):e038707. <https://doi.org/10.1136/bmjopen-2020-038707> PMID: [33408198](https://pubmed.ncbi.nlm.nih.gov/33408198/)
65. Kaat AJ, Rothrock NE, Vrahas MS, O'Toole R V., Buono SK, Zerhusen T, et al. Longitudinal Validation of the PROMIS Physical Function Item Bank in Upper Extremity Trauma. *J Orthop Trauma*. 2017; 31(10):e321–6. <https://doi.org/10.1097/BOT.0000000000000924> PMID: [28938284](https://pubmed.ncbi.nlm.nih.gov/28938284/)
66. Gabbe BJ, Simpson PM, Lyons RA, Ameratunga S, Harrison JE, Derrett S, et al. Association between the number of injuries sustained and 12-month disability outcomes: Evidence from the injury-VIBES study. *PLoS One*. 2014; 9(12):1–15. <https://doi.org/10.1371/journal.pone.0113467> PMID: [25501651](https://pubmed.ncbi.nlm.nih.gov/25501651/)
67. Burd RS, Jensen AR, VanBuren JM, Alvey JS, Richards R, Holubkov R, et al. Long-Term Outcomes after Pediatric Injury: Results of the Assessment of Functional Outcomes and Health-Related Quality of Life after Pediatric Trauma Study. *J Am Coll Surg*. 2021; 233(6):666–675.e2. <https://doi.org/10.1016/j.jamcollsurg.2021.08.693> PMID: [34592405](https://pubmed.ncbi.nlm.nih.gov/34592405/)
68. de Graaf MW, Reininga IHF, Wendt KW, Heineman E, El Moumni M. The Short Musculoskeletal Function Assessment: a study of the reliability, construct validity and responsiveness in patients sustaining trauma. *Clin Rehabil*. 2019; 33(5):923–35. <https://doi.org/10.1177/0269215519828152> PMID: [30722686](https://pubmed.ncbi.nlm.nih.gov/30722686/)
69. Cohen J. Statistical power analysis for the behavioral sciences. Second edi. New York: Lawrence Erlbaum Associates; 1988.

70. Ratner B. The correlation coefficient: Its values range between +1/-1, or do they? *J Targeting, Meas Anal Mark.* 2009; 17(2):139–42.
71. Ponsford J, Hill B, Karamitsios M, Bahar-Fuchs A. Factors influencing outcome after orthopedic trauma. *J Trauma.* 2008; 64(4):1001–9. <https://doi.org/10.1097/TA.0b013e31809fec16> PMID: 18404068
72. Herrera-escobar JP, Apoj M, Weed C, Harlow AF, Sanam S, Rafai A, et al. Association of pain after trauma with long-term functional and mental health outcomes. *J Trauma Acute Care Surg.* 2018; 85(4):773–9. <https://doi.org/10.1097/TA.0000000000002017> PMID: 30020227
73. de Vet HC, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine.* Cambridge University Press, editor. Cambridge; 2018.
74. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med.* 2016; 15(2):155–63. <https://doi.org/10.1016/j.jcm.2016.02.012> PMID: 27330520
75. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986; 1(8476):307–10. PMID: 2868172
76. Yang CM, Wang Y ching, Lee C hua, Chen M hsiang, Hseih CL. A comparison of test–retest reliability and random measurement error of the Barthel Index and modified Barthel Index in patients with chronic stroke. *Disabil Rehabil.* 2022; 44(10):2099–103. <https://doi.org/10.1080/09638288.2020.1814429> PMID: 32903114
77. Kaat AJ, Buckenmaier C III, Cook KF, Rothrock NE, Schalet BD, Gershon RC, et al. The expansion and validation of a new upper extremity item bank for the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Patient-Reported Outcomes.* 2019; 3(69).
78. Hart DL. Assessment of unidimensionality of physical functioning in patients receiving therapy in acute orthopedic outpatient centers. *J Outcome Meas.* 2000; 4(1):413–30. PMID: 11272593
79. Glenny C, Stolee P. Comparing the Functional Independence Measure and the interRAI / MDS for use in the functional assessment of older adults: a review of the literature. *BMC Geriatr.* 2009; 9(52):1–12. <https://doi.org/10.1186/1471-2318-9-52> PMID: 19943969
80. Ravaud JF, Delcey M, Yelnik A. Construct validity of the Functional Independence Measure (FIM): questioning the unidimensionality of the scale and the “value” of FIM scores. *Scand J Rehabil Med.* 1999; 31:31–41. <https://doi.org/10.1080/003655099444704> PMID: 10230001
81. Gunn S, Burgess GH, Maltby J. A Factor Analysis of Functional Independence and Functional Assessment Measure Scores Among Focal and Diffuse Brain Injury Patients: The Importance of Bifactor Models. *Arch Phys Med Rehabil.* 2018; 99(9):1805–10. <https://doi.org/10.1016/j.apmr.2018.04.005> PMID: 29715437
82. Laake K, Laake P, Ranhoff AH, Sveen U, Wyller TB, Bautz-Holter E. The Barthel ADL Index: Factor Structure depends upon the Category of Patient. *Age Ageing.* 1995; 24:393–7. <https://doi.org/10.1093/ageing/24.5.393> PMID: 8669341
83. De Graaf MW, Reininga IHF, Wendt KW, Heineman E, El Mounni M. Structural validity of the Short Musculoskeletal Function Assessment in patients with injuries. *Phys Ther.* 2018; 98(11):955–67. <https://doi.org/10.1093/ptj/pzy098> PMID: 30107487
84. Pretz CR, Kean J, Heinemann AW, Kozlowski AJ, Bode RK, Gebhardt E. A Multidimensional Rasch Analysis of the Functional Independence Measure Based on the National Institute on Disability, Independent Living, and Rehabilitation. *J Neurotrauma.* 2016; 33:1358–62.
85. Hays RD, Spritzer KL, Amtmann D, Lai JS, Dewitt EM, Rothrock N, et al. Upper-extremity and mobility subdomains from the Patient-Reported Outcomes Measurement Information System (PROMIS) adult physical functioning item bank. *Arch Phys Med Rehabil.* 2013; 94(11):2291–6. <https://doi.org/10.1016/j.apmr.2013.05.014> PMID: 23751290
86. Jette AM, Haley SM, Tao W, Ni P, Moed R, Meyers D, et al. Prospective evaluation of the AM-PAC-CAT in outpatient rehabilitation settings. *Phys Ther.* 2007; 87(4):385–98. <https://doi.org/10.2522/ptj.20060121> PMID: 17311888
87. Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): Scale development, measurement properties, and clinical application. *Phys Ther.* 1999; 79(4):371–83.
88. Sepehri A, Sleat GJK, Roffey DM, Broekhuysen HM, O'Brien PJ, Guy P, et al. Responsiveness of the PROMIS physical function measure in orthopaedic trauma patients. *Injury.* 2022; 53(6):2041–6. <https://doi.org/10.1016/j.injury.2022.03.017> PMID: 35300869
89. Dodds TA, Martin DP, Stolov WC, Deyo RA. A validation of the Functional Independence Measurement and its performance among rehabilitation inpatients. *Arch Phys Med Rehabil.* 1993; 74(5):531–6. [https://doi.org/10.1016/0003-9993\(93\)90119-u](https://doi.org/10.1016/0003-9993(93)90119-u) PMID: 8489365

90. Bosse MJ, Mackenzie EJ, Kellam JF, Burgess AR, Webb LX, Swiontkowski M, et al. An analysis of outcomes of reconstruction or amputation of leg-threatening injuries. *N Engl J Med*. 2002; 347(24):1924–31.
91. Kucukdeveci AA, Yavuzer G, Tennant A, Suldur N, Sonel B, Arasil T. Adaptation of the modified Barthel Index for use in physical medicine and rehabilitation in Turkey. *Scand J Rehabil Med*. 2000; 32(2):87–92. PMID: [10853723](https://pubmed.ncbi.nlm.nih.gov/10853723/)
92. Tsuji T, Sonoda S, Domen K, Saitoh E, Liu M, Chino N. ADL structure for stroke patients in Japan based on the Functional Independence Measure. *Am J Phys Med Rehabil*. 1995; 74(6):432–8. <https://doi.org/10.1097/00002060-199511000-00007> PMID: [8534387](https://pubmed.ncbi.nlm.nih.gov/8534387/)
93. Lawton G, Lundgren-Nilsson Å, Biering-s F, Tesio L, Slade A, Penta M, et al. Cross-cultural validity of FIM in spinal cord injury. *Spinal Cord*. 2006; 44:746–52. <https://doi.org/10.1038/sj.sc.3101895> PMID: [16389268](https://pubmed.ncbi.nlm.nih.gov/16389268/)
94. Pransky G, Feuerstein M, Himmelstein J, Katz JN, Vickers-Lahti M. Measuring functional outcomes in work-related upper extremity disorders: Development and validation of the upper extremity function scale. *J Occup Environ Med*. 1997; 39(12):1195–202.
95. Mehta SP, Fulton A, Quach C, Thistle M, Toledo C, Evans NA. Measurement Properties of the Lower Extremity Functional Scale: A Systematic Review. *J Orthop Sports Phys Ther*. 2016; 46(3):200–16. <https://doi.org/10.2519/jospt.2016.6165> PMID: [26813750](https://pubmed.ncbi.nlm.nih.gov/26813750/)
96. Haley SM, Coster WJ, Andres PL, Ludlow LH, Ni Pengsheng, Bond TLY, et al. Activity outcome measurement for postacute care. *Med Care*. 2004; 42(1 Suppl):49–61. <https://doi.org/10.1097/01.mlr.0000103520.43902.6c> PMID: [14707755](https://pubmed.ncbi.nlm.nih.gov/14707755/)
97. Ottenbacher KJ, Hsu Y, Granger CV., Fiedler RC. The reliability of the functional independence measure: A quantitative review. *Arch Phys Med Rehabil*. 1996; 77(12):1226–32.
98. Naghdi S, Ansari NN, Raji P, Shamili A, Amini M, Hasson S, et al. Cross-cultural validation of the Persian version of the Functional Independence Measure for patients with stroke. *Disabil Rehabil*. 2016; 38(3):289–98. <https://doi.org/10.3109/09638288.2015.1036173> PMID: [25885666](https://pubmed.ncbi.nlm.nih.gov/25885666/)
99. Wild H, Stewart BT, LeBoa C, Stave CD, Wren SM. Epidemiology of Injuries Sustained by Civilians and Local Combatants in Contemporary Armed Conflict: An Appeal for a Shared Trauma Registry Among Humanitarian Actors. *World J Surg*. 2020; 44(6):1863–73. <https://doi.org/10.1007/s00268-020-05428-y> PMID: [32100067](https://pubmed.ncbi.nlm.nih.gov/32100067/)
100. Breivik H. Fifty years on the Visual Analogue Scale (VAS) for pain-intensity is still good for acute pain. But multidimensional assessment is needed for chronic pain. *Scand J Pain*. 2016; 11:150–2. <https://doi.org/10.1016/j.sjpain.2016.02.004> PMID: [28850458](https://pubmed.ncbi.nlm.nih.gov/28850458/)
101. Breivik H. How to implement an acute pain service. *Best Pract Res Clin Anaesthesiol*. 2002; 16(4):527–47. <https://doi.org/10.1053/bean.2002.0259> PMID: [12516890](https://pubmed.ncbi.nlm.nih.gov/12516890/)
102. Grzelak S, Bérubé M, Pelet S, Belzile É. Pain Management Strategies After Orthopaedic Trauma: A Mixed-Methods Study with a View to Optimizing Practices. *J Pain Res*. 2022; 15:385–402. <https://doi.org/10.2147/JPR.S342627> PMID: [35177930](https://pubmed.ncbi.nlm.nih.gov/35177930/)
103. Li H, Nyland J, Kuban K, Givens J. Physical therapy needs for patients with physical function injuries post - earthquake disasters: A systematic review of Chinese and Western literature. *Physiother Res Int*. 2018; 23:1–28. <https://doi.org/10.1002/pri.1714> PMID: [29608038](https://pubmed.ncbi.nlm.nih.gov/29608038/)
104. Sullivan GM, Feinn R. Using Effect Size—or Why the P Value Is Not Enough. *J Grad Med Educ*. 2012; (September):279–82.
105. Pathak A, Sharma S, Heinemann AW, Stratford PW, Ribeiro DC, Abbott JH. Development and assessment of a verbal response scale for the Patient-Specific Functional Scale (PSFS) in a low-literacy, non-western population. *Qual Life Res*. 2021; 30(2):613–28. <https://doi.org/10.1007/s11136-020-02640-0> PMID: [32965632](https://pubmed.ncbi.nlm.nih.gov/32965632/)
106. Thong ISK, Jensen MP, Tan G. Original experimental The validity of pain intensity measures: what do the NRS, VAS, VRS, and FPS-R measure? *Scand J Pain*. 2018; 18(1):99–107.
107. Kucukdeveci AA, Tennant A, Grimby G, Franchignoni F. Strategies for assessment and outcome measurement in physical and rehabilitation medicine: An educational review. *J Rehabil Med*. 2011; 43(8):661–72. <https://doi.org/10.2340/16501977-0844> PMID: [21687922](https://pubmed.ncbi.nlm.nih.gov/21687922/)
108. World Health Organization. WHA76.6 Strengthening rehabilitation in health systems. 2023.
109. World Health Organization. Global report on health data systems and capacity, 2020. Geneva; 2020.
110. Koumamba AP, Bisvigou UJ, Ngoungou EB, Diallo G. Health information systems in developing countries: case of African countries. *BMC Med Inform Decis Mak*. 2021; 21(232):1–10. <https://doi.org/10.1186/s12911-021-01597-5> PMID: [34348718](https://pubmed.ncbi.nlm.nih.gov/34348718/)