

Translation, cross-cultural adaptation and validation of Barthel index into Nepali language

Govinda Mani Nepal

govinda@kusms.edu.np

Kathmandu University

Veronica Ntsiea

University of the Witwatersrand

Ranjeeta S Acharya

Kathmandu University

Ole Petter Norvang

Norwegian University of Science and Technology

Ann-Katrin Stensdotter

Norwegian University of Science and Technology

Research Article

Keywords: Activities of Daily Living, Reliability, Responsiveness, Stroke, Validity

Posted Date: September 30th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10996941/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Translation, cross-cultural adaptation and validation of Barthel index into Nepali language

ABSTRACT

Background: Monitoring activities of daily living (ADLs) during stroke rehabilitation is essential for tracking physical recovery. Nepal lacks context-specific stroke-related tools for assessment of ADLs. This study aimed to translate, culturally adapt, and validate the Barthel Index (BI) into Nepali (BI-NP).

Methodology: BI was translated and adapted using Mapi Trust's linguistic validation guidelines for clinical outcome assessments. Psychometric properties of BI-NP were evaluated following COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) criteria. To assess validity, Nepali versions of WHO Disability Assessment Schedule 2.0, Patient Specific Functional Scale, and Timed Up and Go Test were administered to 76 stroke survivors. For intra- and inter-rater reliability and responsiveness, the Nepali Global Rating of Change was used as an external anchor to classify participants into stable and improved groups.

Results: BI was successfully translated and culturally adapted. The minimum detectable change for the total BI score (range 0-100) between test and re-test was 7.24. Cronbach's alpha for the total BI score was 0.9, indicating excellent internal consistency. Intra-rater and inter-rater

24 reliability were 0.999 (95% CI: 0.999–0.999) and 0.997 (95% CI: 0.996–0.998) at test and re-
25 test, respectively. Responsiveness analysis showed an area under the curve of 0.8.

26 **Conclusions:** BI-NP is a valid, reliable, and responsive tool for assessing functional
27 independence in Nepali stroke survivors.

28 **KEYWORDS:** Activities of Daily Living, Reliability, Responsiveness, Stroke, Validity

29

30

INTRODUCTION

Stroke is a worldwide health concern, ranking as the second most common cause of death, and comprising 5.7% of all disability-adjusted life years (DALYs) [1]. In Nepal, stroke is one of the most common causes of death contributing to 3.5% of total DALYs [2]. The stroke-related DALY rate of low-income countries was reported to be 3.7 times higher than that of the high-income countries, emphasizing the considerable impact on developing countries like Nepal [3].

Stroke impairs activities of daily living (ADLs), with mobility loss being a major contributor to reduced independence and social participation [4, 5]. Physiotherapy interventions therefore emphasize restoring mobility and ADLs function [6, 7]. Reliable functional assessments are essential to evaluate performance and monitor therapeutic progress [8]. These measures are central to stroke rehabilitation, guiding clinical decision-making and research by assessing treatment outcomes, rehabilitation eligibility, independence, and disability levels [9-12].

Among the commonly used functional outcome tools to assess a person's level of independence after stroke, the Barthel Index (BI) is a reliable and valid 10-item measure of basic ADLs [13], with total scores ranging from 0 to 100; higher scores indicate greater independence. Its widespread acceptance is attributed to simplicity, cost-effectiveness, ease of use, and efficiency [13, 14]. It has been translated into 32 different languages [15] and has repeatedly shown to be a reliable and valid tool to measure ADLs [16]. At present, there is no standardized or validated Nepali-language tool available to assess ADLs among individuals with stroke in Nepal.

Functional status is commonly evaluated through informal clinical observations or with non-contextual English-language instruments, which are rarely utilized due to language barriers, cost, complexity, and time constraints. The scarcity of rehabilitation centers and shortage of trained professionals further limit the feasibility of complex assessments. Consequently, the current

approach results in inconsistent documentation of patient progress, limited comparability of outcomes across settings, and inadequate evidence to guide individualized rehabilitation planning.

A translated and cross-cultural adapted BI into Nepali language and context may amend this. BI measures basic mobility and self-care tasks including walking, stair climbing, toileting, and feeding, thus offers a practical and standardized assessment tool for rehabilitation professionals, community health workers, and caregivers. Its use can strengthen clinical decision-making, facilitate goal setting, improve continuity of care, and generate outcome data to inform service planning and policy, thereby enhancing the quality and equity of stroke rehabilitation services in Nepal. To ascertain the measurement properties including reliability, validity and responsiveness of a Nepali version of BI, it is imperative to adhere to standard guidelines to establish it as a context-specific outcome measure for stroke [17].

Thus, this study was designed to cross-culturally adapt the BI and translate it into the Nepali language and to assess its measurement properties among Nepali stroke survivors. The aim was to provide a reliable, valid and responsive tool for professionals to assess independence in ADLs among Nepali stroke survivors.

METHODOLOGY

This study comprised 1) translation and cross-cultural adaptation and 2) testing measurement properties of the BI for use in Nepal, hence denoted BI-NP. Translation and cross-cultural adaptation of BI into the Nepali language and context was conducted applying the standard linguistic validation guidance of a clinical outcome assessment (COA) as recommended by the Mapi Trust [18]. Measurement properties of BI-NP were validated in patients with diagnosed

stroke admitted at two tertiary hospitals: 1) Annapurna Neurological Institute and Allied Sciences (ANIAS) located in Kathmandu valley and 2) Dhulikhel Hospital located 30 Kilometers outside Kathmandu. ANIAS is among a few dedicated centers located in urban Nepal providing care after stroke. Dhulikhel hospital is a tertiary care center located in sub-urban Nepal. The sites were selected for inclusion of both urban and sub-urban participants and convenience of data collection. Data was collected from 18th July 2023 to 18th May 2024 by using non-probability convenience sampling.

Participants

Translation and cross-cultural adaptation

Forward translation was performed by two native Nepalese speakers proficient in English. Back translation was performed by a professional translator with English as native language and a native Nepali speaker. Comparison between the backward version and source instrument was done by the primary researcher.

Testing measurement properties of the BI-NP

Responsiveness, reliability and validity were tested by two trained physiotherapy interns. Participants were Nepali speaking patients aged >18 years of age with ischemic or haemorrhagic stroke in acute or sub-acute phase. Exclusion criteria were speech abnormalities, cognitive deficits or psychiatric illness (assessed with Nepali version of Rowland Universal Dementia Assessment Scale (N-RUDAS) score ≤ 22), history of transient ischemic attacks, and other conditions impacting ADLs.

Permission and ethics

The study was conducted in line with the Helsinki declaration [19]. Permission for translating BI into the Nepali version was obtained from copyright holder ‘Mapi Research Trust’ (86842). Written consent was obtained from all participants prior to data collection.

Procedure for translation and cross-cultural adaptation

The translation and cross-cultural adaptation followed five steps (Figure 1). The pre-final scale was pre-tested with seven stroke survivors, requiring no modifications. Final proofreading addressed minor grammatical and formatting issues before submission for appraisal.

The primary researcher proofread the final version which, along with the reports of the translation and cross-cultural adaptation process, was sent to the Mapi Research Trust for appraisal [20]. No further modifications were needed, and the final version of BI-NP was generated (Appendix A).

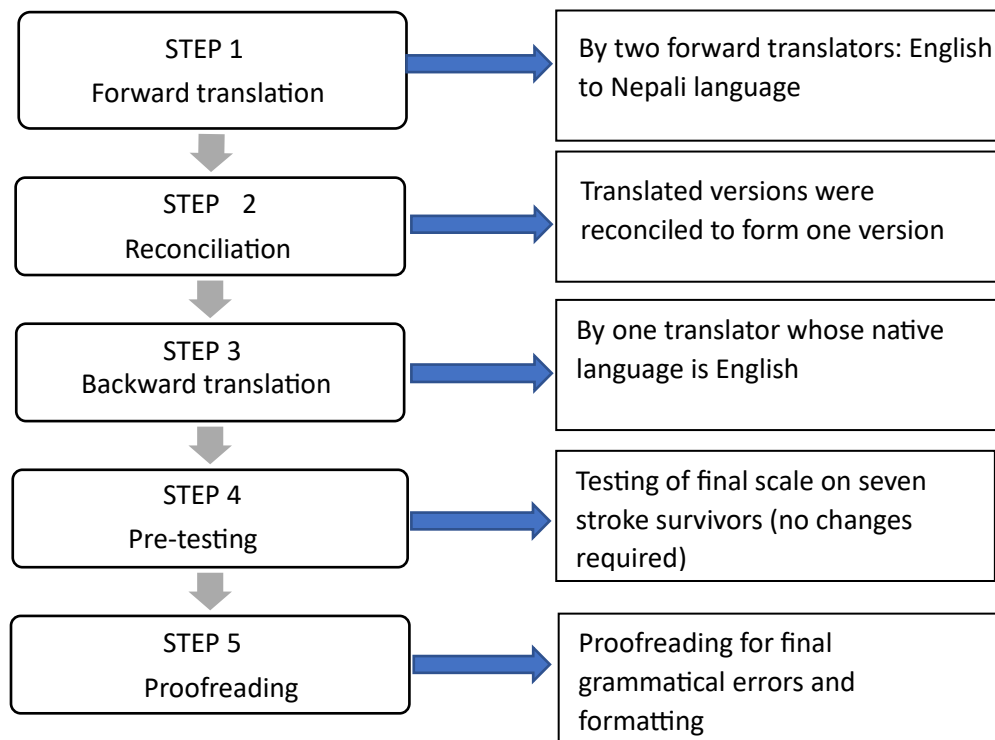


Figure 1: Flow diagram showing steps of translation and cross-cultural adaptation (recommended by the Mapi Trust [20])

Procedures for testing measurement properties

BI-NP was evaluated for validity, reliability and responsiveness according to the COSMIN criteria [21]. For inter-rater reliability and content validity, two blinded physiotherapists independently administered the BI-NP, the Nepali versions of WHO Disability Assessment Schedule 2.0 (WHODAS 2.0-NP) [22], Patient Specific Functional Scale (PSFS-NP) [23], and Time Up and Go (TUG-NP) [24] to stroke survivors (n=76). The added value of these tools was considered the best option for evaluating content validity as the collective ADLs items incorporated in these tools aligns with those in the BI-NP (3 of 12 items in WHODAS 2.0-NP, 5 in PSFS-NP and 1 in TUG-NP). For intra-rater reliability and responsiveness, Global Rating of Change (GROC-NP) [25] was

added to monitor potential change in the patient's functional status between tests across two consecutive days to balance the stability of patients and recall bias.

Responsiveness was tested at baseline and follow up at two weeks (n=43) [26] to detect change over time in the construct, according to COSMIN [21] and evaluated as recommended by deVet [27]. To determine sensitivity and specificity, GROC-NP was used as the external anchor for the assessment of functional independence using BI-NP [25, 28, 29]. For construct validity we hypothesized that BI-NP would correlate positively with 1) PSFS-NP, and 2) negatively with WHODAS 2.0-NP and 3) TUG-NP. BI-NP would be considered to have acceptable construct validity if at least two out of three priori hypotheses were met [21]. Structural validity was assessed with factor analysis to examine whether the scores reflected the underlying dimensionality of the test.

Statistical analyses

The statistical analyses were conducted using the Statistical Package for Social Science (SPSS) Version 25 (IBM Corp., Armonk, NY).

Validity

Normality was tested using Kolmogorov–Smirnov normality test ($p \geq 0.05$). For ordinal, non-parametric tests Spearman's correlation coefficient (rho) was used to determine the magnitude of the correlation between the BI-NP, WHODAS 2.0-NP, PSFS-NP, and TUG-NP. The correlation coefficients were interpreted using Cohen's thresholds ($> \pm 0.5$ strong, ± 0.5 – 0.3 moderate, ± 0.3 – 0.1 weak, $< \pm 0.1$ insubstantial) [30, 31]. The structural validity was assessed by exploratory factor analysis (EFA) with principal axis factoring extraction. Sampling adequacy was examined with Kaiser-Meyer Olkin (KMO) for suitability of the data for factor analysis and confirmed with

Bartlett's test of sphericity. The number of significant factors was determined by evaluation of the eigenvalue of each factor (criterion >1).

Reliability

Cronbach's alpha was used to assess internal consistency to measure the homogeneity of the scale and the interrelatedness of the items [32]. Intra- and inter-rater reliability was estimated using two-way mixed effects models with absolute agreement and the resulting intraclass correlation coefficient (ICC) values demonstrating good (>0.70) to excellent (>0.75) agreement [33].

Measurement error

Minimal detectable change (MDC) was calculated using the formula $MDC = z \times \sqrt{2} \times SEM$, where $z = 1.96$ (score for estimating a 95% confidence interval), $\sqrt{2}$ indicates the two BI-NP measurements and SEM is the standard error of measurements ($SEM = SD (1 - ICC)^{1/2}$, where SD is the standard deviation) for the mean change of BI-NP score from baseline to follow-up measurement, and ICC is the intraclass correlation coefficient of the stable group.

Responsiveness

Responsiveness was assessed using a Receiver operator characteristic (ROC) curve for accuracy of the classification between changes in BI-NP scores and GROC-NP. To determine sensitivity and specificity, the area under curve (AUC) of 0.7 or more was considered acceptable for the ability of BI-NP to detect a change [25, 34, 35]. The minimal important change (MIC) of BI-NP was also extracted from the ROC curves.

Floor and ceiling effects

We considered the presence of floor and ceiling effects if >15% of the respondents achieved the lowest or highest possible BI-NP total scores, respectively [36].

RESULTS

Translation and cross-cultural adaptation

The process of translating and culturally adapting BI for Nepali culture provided several challenges and prompted adjustments in three main areas (Appendix B). The major changes that were made are narrated in Table 1.

Table 1: Examples of forward and backward translation during synthesis and adaptation of the BI-NP

Original BI English	Translation to Nepali	Back translation to English	Agreed Nepali
Bed	Wochhyan (Nepali) ¹ Khaat (Nepali) ²	Bed	Khaat
Razor	Daari katne machine	Shaving machine (English) ¹ Razor (English) ²	Daari katne machine
Doing personal toilet	Bektigaat sarsafai garne (Nepali) ¹ Bektigaat sringaar garne (Nepali) ²	Maintaining personal hygiene	Bektigaat sarsafai garne
Item	Ekaai Bundha	Components	Bundha [37]
Dichotomous	dui bhagma bivajan garieko	Divided into two parts	dui bhagma bivajan garieko
Flight of stairs	Ek tala varyang	One floor	Ek tala varyang

[A] Semantic equivalence: Several linguistic differences and grammatical preferences for the same statement were noted among the translators. Linguistic changes were made to resonate with the best translation available during the review.

[B] Conceptual equivalence: Adaptations were made to reflect Nepali cultural practices. Culturally appropriate eating tasks like "breaking the food into pieces" and "mixing the food with hands" were added to item 1 ("Feeding"), acknowledging that many Nepali people eat with their hands. Likewise, the term "toilet paper" was supplemented with "Paani" (water), to align with Nepali cultural practices in items 3 and 4. For item 5, different terms for bathing techniques like

"Baltibata khanaune" (pouring from bucket), "Dharabata jieuma thapera nuhaune" (pouring from the tap), and "Jalo"(sponge) were added along with the terms "bathtub", "shower", or "sponge" presented in the English version. For item 6, "Walking on level surface", the measurement unit "Yard" was converted to "feet", the measurement unit used in Nepal.

[C] Idiomatic equivalence: We could not translate the term 'girdle' in item 8 since it had no distinct meaning in Nepali and was retained as it is. On items 9 and 10, we did not find the relevant literal Nepali word for "accidents" and hence, adapted the term "chuhiyeko" from the Nepali version of Spinal Cord Independent Measure (N-SCIM) [38].

Measurement properties

A total of 76 participants with stroke were included in the study for reliability and validity and a sub-sample of 43 at follow up were evaluated on two occasions (with an interval of 2 weeks) for responsiveness. The demographic and clinical characteristics of participants are presented in Table 2.

Table 2: Demographic and clinical characteristics of participants (N=76)

Variable	Mean $\pm$ SD	N (%)
Age	62.1 $\pm$ 14.093	76
Sex		
Male		51 (67.1)
Female		25 (32.9)
Occupation		
Farmer		25 (32.9)
Housewife		19 (25.0)
Others		32 (42.1)
Marital status		
Married		72 (94.7)
Unmarried		4 (5.3)
Education		
Grade 1-5		12 (15.8)
Grade 6-9		5 (6.6)
Grade 10-12		14 (18.4)
Bachelor's/Master		5 (6.6)

Informal	9 (11.2)
Cannot read and write	30 (39.5)
Hand Dominance	203
Right	74 (97.4)
Left	2 (2.6)
Time from Stroke to inclusion	205
Acute (first two weeks of onset)	56 (73.7)
Sub-acute (3-11 weeks post-stroke)	20 (26.3)
Stroke history	207
First	71 (93.4)
Recurrent	5 (6.5)
Classification of stroke	208
Ischemic	39 (51.2)
Total anterior circulation infarct	14 (35.9)
Partial anterior circulation infarct	6 (15.4)
Posterior circulation Infarct	13 (33.3)
Lacunar infarct	6 (15.4)
Hemorrhagic	37 (48.2)
Intracerebral hemorrhage	11 (29.7)
Basal ganglia bleed	14 (37.8)
Thalamic bleed	12 (32.4)

Abbreviations: **SD: standard deviation**

215

216 The internal consistency, intra-rater reliability and inter-rater reliability of the BI-NP are shown in
217 Table 3, indicating excellent reliability. The median GROC score for intra-rater reliability was
218 found to be 5, denoting no change in functional level within two consecutive assessment days.
219 SEM was 2.61 points and MDC at the 95% confidence level was calculated as 7.24 points.
220 For construct validity (Table 3), a limited number of ADLs items in each of the correlated tools
221 were suitable for comparison with BI-NP (Appendix C).

222

223 **Table 3: Reliability and validity measure of BI-NP (N=76)**

Measurement properties	BI-NP	Interpretation
Internal consistency (Cronbach's alpha)	0.90	Excellent
Intra-rater reliability ICC (95%CI)	0.99 (0.999-0.999)	Excellent
Inter-rater reliability ICC (95%CI)	0.99 (0.996-0.998)	Excellent
SEM	2.61	
MDC	7.24	
Construct validity		
BI-NP & WHODAS 2.0-NP*		
Bathing self/Washing your whole body	r= -0.601	Moderate to strong negative correlation
Walking on level surface/Walking a long distance	r= -0.451	
Dressing and undressing/Getting dressed	r= -0.839	
BI-NP & PSFS-NP*		

Feeding	r= 0.878	
Moving from wheelchair to bed and return	r=0.937	Strong positive correlation
Getting on and off toilet	r=0.897	
Ascending and descending stairs	r= 0.753	
Continence of bladder	r=0.785	
BI-NP & TUG-NP*		Strong negative correlation
Walking on level surface & walking 3 meters	r= -0.779	

Abbreviations: BI-NP Nepali version of Barthel Index, CI confidence interval, MDC minimum detectable change, PSFS-NP patient specific functional scale Nepali, SEM standard error of measurement, TUG-NP time up and go test Nepali, WHODAS 2.0-NP WHO disability assessment schedule 2.0 Nepali
*level of significance (p)<0.01

The structural validity (factor analysis) of BI-NP with 10 items revealed multifactorial structure (Kaiser-Meier-Olkin 0.883) with two factors emerging with Eigen values >1 which accounted for 70.2% of the total variance. The first factor included eight items and accounted for 58.1% and the second factor included two items accounted for 12.1% with loading > 0.6 (Table 4).

Table 4: Two factor structure of BI-NP

Item Number	Barthel Index Items (Nepali)	Barthel Index Items (English)	Factor loading	
			1	2
1	खाना खाने	Feeding	0.723	0.173
2	क्लिलचियरबाट ओछ्यानमा सर्ने र फर्किने	Moving from wheelchair to bed and return	0.795	0.097
3	व्यक्तिगत सरसफाई गर्ने	Doing personal toilet	0.803	0.012
4	शौचालयको प्रयोग गर्ने	Getting on and off toilet	0.840	0.118
5	आफै नुहाउने	Bathing self	0.649	-0.092
6	समतल सतहमा हिँड्ने	Walking on level surface	0.896	-0.089
7	भर्याङ चढ्ने वा ओर्ल्ने	Ascending and descending stairs	0.809	0.010
8	लुगा लगाउने र खोल्ने	Dressing and undressing	0.895	-0.044
9	दिसाको नियन्त्रण गर्ने	Continence of bowel	-0.091	0.962*
10	पिसावको नियन्त्रण गर्ने	Continence of bladder	0.212	0.780*

Extraction Method: Principal Component Analysis; Rotation Method: Oblimin with Kaiser Normalization; *Emergence of second factor with loading > 0.6

Abbreviations: BI-NP Nepali version of Barthel index

The responsiveness of the BI-NP based on the GROC-NP scores, as indicated by an AUC value, was 0.8 (Figure 2). The identified optimal threshold of the Youden Index, MIC of 7.50 on the BI-

NP change scale, balanced sensitivity and specificity effectively. The sensitivity of 0.84, accurately identified 84% of truly improved cases, while the specificity of 0.81, indicated the correct identification of 81% of truly not improved cases (Figure 3).

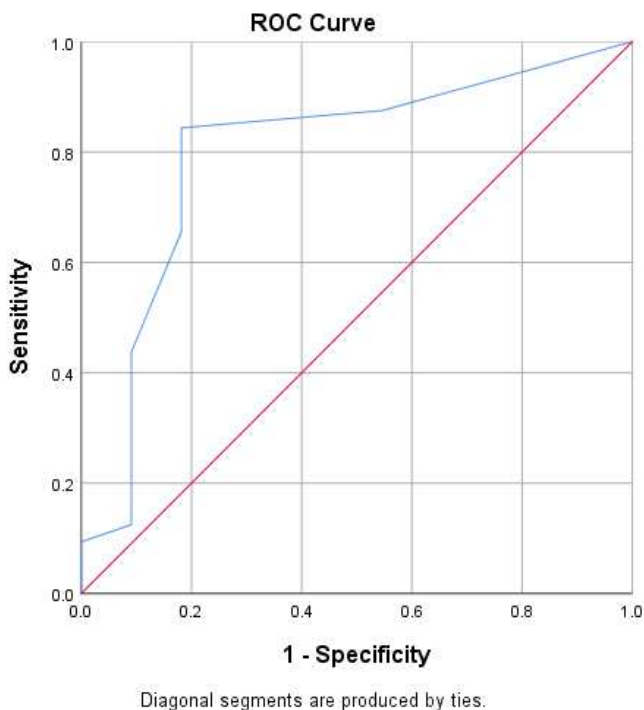


Figure 2: ROC curve between stable and improved groups

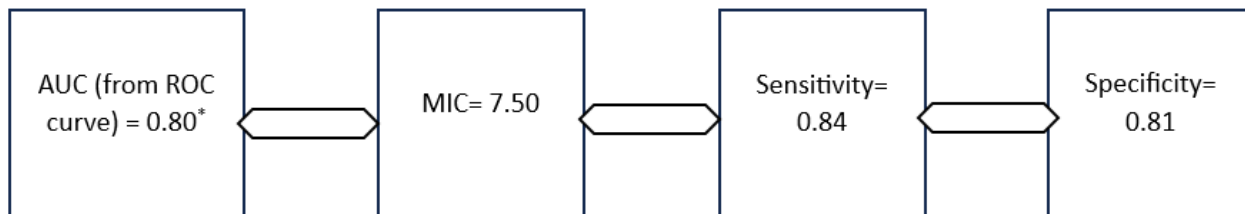


Figure 3: Responsiveness measure of BI-NP (N=43)

Abbreviations: AUC area under curve, BI-NP Nepali version of Barthel index, MIC Minimal important change, ROC curve receiver operating characteristic curve

*level of significance (p)<0.05

Floor and ceiling effects

No floor or ceiling effects were detected at baseline, with only 3.9% of participants obtaining the lowest possible scores, and 1.3% obtaining the highest possible scores.

DISCUSSION

The 10-item BI was successfully translated and cross culturally adapted into Nepali. Several adaptations were made in terms of basic translation of words/phrases, context and culture to address the linguistic diversity across the Nepali community. The measurement properties of the BI-NP among stroke survivors exhibited good to excellent validity, reliability and acceptable responsiveness.

Translation and cross-cultural adaptation

BI has been translated and cross culturally adapted into several languages, with varying degrees of modification required [39-50]. During translation into Nepali, language differences emerged during forward translation and synthesis, consistent with findings by Yogendra et al. [51], reflecting the linguistic diversity of the Nepali community. Expert review further identified ambiguities between the original English and the Nepali versions of several items, requiring contextual adaptations to ensure conceptual equivalence and cultural relevance.

The adaptations reflected common daily practices and living environments in Nepal rather than direct linguistic substitution. Nepal represents a mix of urban and rural settings, where household environments and functional routines vary considerably. Activities of daily living may be performed differently from the assumptions embedded in the original BI, particularly for eating, personal hygiene, bathing, mobility, and stair use. Therefore, modifications were introduced to

retain the intended construct while improving comprehensibility and applicability in the local context.

Major revisions were made to items related to feeding, personal toilet, and bathing, while minor linguistic refinements were applied across the remaining items. For feeding, expressions describing food handling were incorporated to reflect common eating practices, similar to adaptations reported in Turkish and Chinese versions. [44, 45]. For personal hygiene, terms were modified to capture locally understood concepts and include commonly used grooming equipment. Bathing-related items were adapted to reflect the diverse bathing practices in Nepal, where alternatives to shower or bathtub bathing, such as bucket bathing and washing under running tap water, remain common. Comparable contextual modifications have also been reported in the Chinese adaptation [45].

Environmental adaptations were also necessary. In the mobility item, “50 yards” was converted to “150 feet” to align with measurement units commonly understood in Nepal, similar to adaptations made in Italian and Nepali SPADI versions [52, 53]. Likewise, the concept of “one flight of stairs” was adapted because household structures in Nepal vary substantially, and stair use is more meaningfully interpreted in relation to moving between floors rather than counting stair segments. Similar contextual considerations have been described in the Chinese version [45]. Collectively, these findings emphasize that cross-cultural adaptation requires preserving conceptual meaning while ensuring relevance to local language, environment, and daily life.

Validity

The three hypotheses about the construct (convergent) validity were supported by strong correlations with WHODAS 2.0, PSFS-NP and TUG-NP items [22, 23, 54]. We have been unable

to find other studies that have investigated construct validity; thus, our results cannot be compared to others. Only 8 out of 10 ADL items in the selected tools matched the items of BI-NP and therefore, only 8 items of the BI-NP were validated. Some of the selected items were still not perfect matches. ‘Walking a long distance such as 1km’ from WHODAS 2.0-NP correlated highly with the item ‘Walking on level surface’ from BI-NP, but being able to walk on a level surface does not say anything about distance. The strength of correlations was very different for WHODAS (-0.451) and TUG (-0.779), as the distance was 1 kilometer or equivalent in WHODAS and only 3 meters in TUG. The participants are likely to rate more difficulty in walking a long distance compared to 3 meters.

Structural Validity (factor analysis)

The EFA in this study revealed a two factor structure solution for the Nepali stroke population similar to the studies conducted in China [55] and Iran [56]. The dimensionality of BI has often been debated, as several studies describe a one-dimensionality nature of BI [57-59], while others have identified a multidimensionality nature [52, 60, 61]. From a conceptual standpoint, functional executive disability (ADLs) and incontinence are two latent categories of functional dependence in stroke populations. This concept has been debated in some literatures where bladder and bowel were considered to be different constructs, suggesting that there is not always a strong correlation between these items and physical dependency or disability [55, 62-64]. The International Classification of Functioning, Disability and Health of the World Health Organization has also defined incontinence as a physiological function impairment rather than a mobility restriction [65]. Factor one obviously describes basic ADLs functions and their relationship with limb mobility whereas factor two, involving continence, denotes a pure physiological function. From nursing

care in a Nepali clinical setting, it is evident that diapers and urinary catheters are used for acute stroke patients, irrespective of the patient's actual need. This might have influenced factor two as a different construct in our study.

Reliability

The alpha coefficient, regarded as the most important index of test reliability [66], revealed excellent internal consistency for BI-NP, coherent with the original version [67] and most previous studies on BI [68, 69]. The high interrelatedness among items indicates that the adapted scale adequately reflects the original English version [70], with a Cronbach's alpha of 0.90, well above the 0.70 threshold recommended for new questionnaires [71]. The Italian version, assessed in ischemic stroke patients through bedside observations, yielded a higher coefficient of 0.94 [52]. Unlike our findings, it showed strong correlations across all items. In Nepal, caregiver assistance with bathing and toileting may explain the lower correlations for 'Bathing self' and continence items. The two-factor structure of BI-NP may also contribute to these differences.

Across diverse contexts, BI has demonstrated strong reliability, with Cronbach's alpha ranging from 0.87 in Dutch spinal cord injury patients to 0.93 in Iranian elderly outpatients [56, 72]. Studies from Japan, Turkey, and China further confirm internal consistency (0.93, 0.88, and 0.92 respectively), underscoring its utility across populations [55, 63, 73].

BI-NP also showed excellent intra- and inter-rater reliability, with ICCs exceeding 0.70 [71], aligning with prior studies [52, 74-76]. A systematic review highlighted BI's outstanding inter-rater reliability across countries, assessor training levels, and delivery modes (face-to-face, phone, video, postal) [77]. These recommendations give an opportunity for a low resource setting like Nepal to utilize the translated tool across various geographical settings, architectural structures,

phone calls as a mode of delivery and even utilizing the caregivers with a modest training on scaling and scoring.

Responsiveness

The responsiveness was calculated with 43 participants only (N=76 on day 1 and N=43 on day 15) as approximately half of them did not come for the follow up. Studies have reported that noncompliance to follow up is a major issues in Nepal healthcare primarily because of lack of awareness and geographical challenges [78]. BI-NP demonstrated adequate capacity to detect clinically relevant changes over time, as indicated by an AUC of 0.80 suggesting a promising level of discriminatory power of BI-NP in distinguishing between patients experiencing improvement and those not experiencing improvement. The result of our study was similar to the Italian culturally adapted version (AUC 0.72), demonstrating its ability to effectively detect clinically significant improvements during hospital rehabilitation [61]. Another study revealed that the BI exhibited good discriminative properties at days 2, 5, and 9 post-stroke, with adequate AUC values of 0.78, 0.83, and 0.84 respectively [79]. The AUC value as noted on day 9 was higher than what we observed in our study (day 15). The patients in our study were discharged to home after their acute management (approximately 5 days). We assume that these patients never received an optimal rehabilitation in their home settings which could explain why the AUC values were lower. Other possible reasons could be the items like ‘Bathing self’, where patients would never have had an opportunity to self-perform the task to demonstrate clinically meaningful changes overtime. Noteworthy is that the variations in responsiveness analysis, such as the inadequate AUC of only 0.65, have also been observed in a study from USA [4]. These discrepancies may result from differences in study populations, settings, methodologies, exercise intensity and dosage, or cultural

adaptations, highlighting the importance of continued research efforts to validate and refine responsiveness assessments of the BI-NP across diverse clinical settings and patient populations.

Strengths of the study

We translated BI using standardized guidelines with the help of expert translators to ensure accurate and reliable translation. Different measurement properties are assessed in this single study. The strength and quality of our study were further enhanced by employing the Nepali version of GROEC for monitoring the progression of stroke patients' functional state during the intra-rater reliability assessment making it more standardized.

Limitations of the study

The limited number of instruments adapted into the Nepali language for functional assessment in stroke survivors led to a smaller number of choices for testing construct validity of BI-NP. The relatively small sample size ($n = 76$; responsiveness subgroup $n = 43$) falls short of the COSMIN guideline recommendation of at least 100 participants for analyses such as internal consistency and structural validity, which may limit the reliability of the exploratory factor analysis findings. Additionally, the absence of a widely recognized comparator, such as the Functional Independence Measure, reduces the strength and interpretability of the construct validity and responsiveness results.

CONCLUSION

The BI-NP was found to be reliable, valid, and responsive for measuring functional independence in Nepali stroke patients. The culturally adapted Barthel Index in Nepali (BI-NP) has the potential to serve as a standardized tool for Nepal's health system for assessing functional independence

among stroke survivors. Its routine use can support clinical decision-making by guiding rehabilitation planning, monitoring patient progress, and informing timely discharge or referral. By identifying individuals with greater functional limitations, the BI-NP can help prioritize rehabilitation services and support evidence-based allocation of limited resources. Furthermore, its simplicity and ease of administration make it feasible for use in community-based care and telerehabilitation, extending its reach to underserved and rural populations. Integrating BI-NP into clinical and public health practice could therefore enhance the quality, equity, and efficiency of stroke rehabilitation services across Nepal. Though Nepali is an official language, there are 124 languages recognized within the country with 142 distinct caste and ethnic groups [80] and thus the external validity of BI-NP should be interpreted cautiously. Further research is needed to ensure its utility across diverse populations and settings.

List of abbreviations

ADLs – Activities of Daily Living
ANIAS – Annapurna Neurological Institute and Allied Sciences
AUC – Area Under the Curve
BI – Barthel Index
BI-NP – Nepali Version of the Barthel Index
CI – Confidence Interval
COA – Clinical Outcome Assessment
COSMIN – COnsensus-based Standards for the selection of health Measurement INstruments
DALYs – Disability-Adjusted Life Years
EFA – Exploratory Factor Analysis
GROC-NP – Nepali Version of the Global Rating of Change
ICC – Intraclass Correlation Coefficient
IRC – Institutional Review Committee
KMO – Kaiser–Meyer–Olkin

418 KUSMS – Kathmandu University School of Medical Sciences
419 MDC – Minimal Detectable Change
420 MIC – Minimal Important Change
421 NHRC – Nepal Health Research Council
422 N-RUDAS – Nepali Version of the Rowland Universal Dementia Assessment Scale
423 N-SCIM – Nepali Version of the Spinal Cord Independence Measure
424 PSFS-NP – Nepali Version of the Patient-Specific Functional Scale
425 ROC – Receiver Operating Characteristic
426 SD – Standard Deviation
427 SEM – Standard Error of Measurement
428 SPSS – Statistical Package for the Social Sciences
429 TUG-NP – Nepali Version of the Timed Up and Go Test
430 USA – United States of America
431 WHODAS – World Health Organization Disability Assessment Schedule

432

433 **Declaration of generative AI and AI-assisted technologies in the manuscript preparation**
434 **process**

435 During the preparation of this work the author(s) used ChatGPT, an AI language model developed
436 by OpenAI, to assist with language refinement, grammar correction, and improvement of
437 manuscript readability. No AI tool was used for data analysis, interpretation, or generation of
438 scientific content. All outputs were reviewed and revised by the authors, who take full
439 responsibility for the manuscript's content.

440

441 **REFERENCES**

- 442 1. Feigin, V.L., et al., *World Stroke Organization (WSO): Global Stroke Fact Sheet 2022*. International
443 Journal of Stroke, 2022. **17**(1): p. 18-29.
444 2. Pyakurel, M.B.S., Joshi B, Koju RP, Shrestha A, “*Burden of Stroke in Nepal: Findings from Global*
445 *Burden of Disease Dataset 2017.*”. Kathmandu University Medical Journal (KUMJ) 2021. **19,75**: p.
446 11-17.

3. *Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019.* Lancet Neurol, 2021. **20**(10): p. 795-820.
4. Wallace, D., P.W. Duncan, and S.M. Lai, *Comparison of the responsiveness of the Barthel Index and the motor component of the Functional Independence Measure in stroke: the impact of using different methods for measuring responsiveness.* J Clin Epidemiol, 2002. **55**(9): p. 922-8.
5. van de Port, I.G., et al., *Susceptibility to deterioration of mobility long-term after stroke: a prospective cohort study.* Stroke, 2006. **37**(1): p. 167-71.
6. Pollock, A., et al., *Physical rehabilitation approaches for the recovery of function and mobility following stroke.* Cochrane Database Syst Rev, 2014. **2014**(4): p. Cd001920.
7. van der Putten, J.J., et al., *Measuring change in disability after inpatient rehabilitation: comparison of the responsiveness of the Barthel index and the Functional Independence Measure.* J Neurol Neurosurg Psychiatry, 1999. **66**(4): p. 480-4.
8. Hsueh, I.P., et al., *Comparison of psychometric properties of three mobility measures for patients with stroke.* Stroke, 2003. **34**(7): p. 1741-5.
9. Dodds, T.A., et al., *A validation of the functional independence measurement and its performance among rehabilitation inpatients.* Arch Phys Med Rehabil, 1993. **74**(5): p. 531-6.
10. Taylor-Rowan, M., et al., *Functional Assessment for Acute Stroke Trials: Properties, Analysis, and Application.* Front Neurol, 2018. **9**: p. 191.
11. Opara, J., et al., *Is the Barthel scale still useful 50 years after its first publication?* 2015. **19**: p. 34-37.
12. Rahayu, U.B., et al., *Effectiveness of physiotherapy interventions in brain plasticity, balance and functional ability in stroke survivors: A randomized controlled trial.* NeuroRehabilitation, 2020. **47**(4): p. 463-470.
13. Quinn, T.J., P. Langhorne, and D.J. Stott, *Barthel Index for Stroke Trials.* Stroke, 2011. **42**(4): p. 1146-1151.
14. Duffy, L., et al., *Reliability (inter-rater agreement) of the Barthel Index for assessment of stroke survivors: systematic review and meta-analysis.* Stroke, 2013. **44**(2): p. 462-468.
15. Mahoney, F.I. and D.W. Barthel, *Functional evaluation: the Barthel Index.* Maryland State Med Journal, 1965. **14**: p. 56-61.
16. Oveisgharan, S., et al., *Barthel Index in a Middle-East Country: Translation, Validity and Reliability.* Cerebrovascular diseases (Basel, Switzerland), 2006. **22**: p. 350-4.
17. Sharrack, B., et al., *The psychometric properties of clinical rating scales used in multiple sclerosis.* Brain, 1999. **122**(1): p. 141-159.
18. Trust, M.R., *Linguistic Validation Guidance of a Clinical Outcome Assessment (COA).* 2016.
19. Association, W.M., *World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects.* JAMA, 2013. **310**(20): p. 2191-2194.
20. Mapi Research Trust. *Translation and linguistic validation.* 2023; Available from: <https://eprovide.mapi-trust.org/instruments/barthel-index>.
21. Mokkink, L.B., et al., *The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content.* BMC Med Res Methodol, 2010. **10**: p. 22.
22. Risal, A., et al., *Adapting World Health Organization Disability Assessment Schedule 2.0 for Nepal.* BMC Psychol, 2021. **9**(1): p. 45.
23. Sharma, S., J. Palanchoke, and J.H. Abbott, *Cross-cultural Adaptation and Validation of the Nepali Translation of the Patient-Specific Functional Scale.* J Orthop Sports Phys Ther, 2018. **48**(8): p. 659-664.
24. Nepal, G.M., M. Basaula, and S. Sharma, *Inter-rater reliability of Timed Up and Go test in older adults measured by physiotherapist and caregivers.* European Journal of Physiotherapy, 2020. **22**(6): p. 325-331.

25. Sharma, S., et al., *Translation, cross-cultural adaptation and psychometric properties of the Nepali versions of numerical pain rating scale and global rating of change*. Health Qual Life Outcomes, 2017. **15**(1): p. 236.
26. Kamper, S.J., C.G. Maher, and G. Mackay, *Global rating of change scales: a review of strengths and weaknesses and considerations for design*. J Man Manip Ther, 2009. **17**(3): p. 163-70.
27. De Vet, H., et al., *Measurement in Medicine: A Practical Guide*. Measurement in Medicine: A Practical Guide, 2011: p. 1-338.
28. Kc, S., et al., *Nepali translation, cross-cultural adaptation and measurement properties of the Shoulder Pain and Disability Index (SPADI)*. J Orthop Surg Res, 2019. **14**(1): p. 284.
29. Kc, S., et al., *Cross-cultural adaptation and measurement properties of the Nepali version of the DASH (disability of arm, shoulder and hand) in patients with shoulder pain*. Health Qual Life Outcomes, 2019. **17**(1): p. 51.
30. Thompson, B., *Effect sizes and confidence intervals for effect sizes*. Psychology in the Schools, 2007. **44**: p. 423-432.
31. Cohen, J., *Set correlation and contingency tables*. Applied psychological measurement, 1988. **12**(4): p. 425-434.
32. Cronbach, L.J., *Coefficient alpha and the internal structure of tests*. Psychometrika, 1951. **16**(3): p. 297-334.
33. Koo, T.K. and M.Y. Li, *A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research*. J Chiropr Med, 2016. **15**(2): p. 155-63.
34. Terwee, C.B., et al., *Quality criteria were proposed for measurement properties of health status questionnaires*. J Clin Epidemiol, 2007. **60**(1): p. 34-42.
35. De Vet, H.C., et al., *Measurement in medicine: a practical guide*. 2011: Cambridge university press.
36. Terwee, C.B., et al., *Quality criteria were proposed for measurement properties of health status questionnaires*. Journal of clinical epidemiology, 2007. **60**(1): p. 34-42.
37. Nepal, G.M., A. Shrestha, and R. Acharya, *Translation and cross-cultural adaptation of the Nepali version of the Rowland universal dementia assessment scale (RUDAS)*. Journal of Patient-Reported Outcomes, 2019. **3**(1): p. 38.
38. Khatri, P., S. Prasertsukdee, and J. Suttiwong, *Reliability of the Nepali Version of the Spinal Cord Independence Measure Self-Report*. Rehabil Res Pract, 2022. **2022**: p. 9983464.
39. Sangha, H., et al., *A comparison of the Barthel Index and the Functional Independence Measure as outcome measures in stroke rehabilitation: patterns of disability scale usage in clinical trials*. International Journal of Rehabilitation Research, 2005. **28**(2): p. 135-139.
40. Hormozi, S., et al., *Iranian version of barthel index: validity and reliability in outpatients' elderly*. International journal of preventive medicine, 2019. **10**.
41. Jung, H.Y., et al., *Development of the Korean version of Modified Barthel Index (K-MBI): multi-center study for subjects with stroke*. Journal of the Korean Academy of Rehabilitation Medicine, 2007. **31**(3): p. 283-297.
42. Cabañero-Martínez, M.J., et al., *The Spanish versions of the Barthel index (BI) and the Katz index (KI) of activities of daily living (ADL): A structured review*. Archives of gerontology and geriatrics, 2009. **49**(1): p. e77-e84.
43. Laohaprasitiporn, P., A. Jarusriwanna, and A. Unnanuntana, *Validity and reliability of the Thai version of the Barthel index for elderly patients with femoral neck fracture*. Journal of the Medical Association of Thailand, 2017. **100**(5): p. 539-48.
44. Küçükdeveci, A.A., et al., *Adaptation of the modified Barthel Index for use in physical medicine and rehabilitation in Turkey*. Scandinavian journal of rehabilitation medicine, 2000. **32**(2): p. 87-92.
45. Leung, S.O., C.C. Chan, and S. Shah, *Development of a Chinese version of the Modified Barthel Index—validity and reliability*. Clinical rehabilitation, 2007. **21**(10): p. 912-922.

46. Mustafa, K., et al., *Translation and Validation of Barthel Index in Urdu Language for Stroke Patients*. Pakistan Journal of Medical & Health Sciences, 2022. **16**(03): p. 163-163.
47. Lorica, J., S. Rosdewi, and M. Solon, *Construct validity Indonesian version of Barthel index for post stroke*. Indian Journal of Public Health, 2020. **11**(03): p. 1997.
48. Heuschmann, P., et al., *The reliability of the german version of the barthel-index and the development of a postal and telephone version for the application on stroke patients*. Fortschritte der Neurologie-psychiatrie, 2005. **73**(2): p. 74-82.
49. Cincura, C., et al., *Validation of the National Institutes of Health Stroke Scale, modified Rankin Scale and Barthel Index in Brazil: the role of cultural adaptation and structured interviewing*. Cerebrovascular Diseases, 2009. **27**(2): p. 119-122.
50. Ohura, T., et al., *Assessment of the validity and internal consistency of a performance evaluation tool based on the Japanese version of the modified barthel index for elderly people living at home*. Journal of Physical Therapy Science, 2014. **26**(12): p. 1971-1974.
51. Yadava, Y.P. *Linguistic diversity in Nepal: Perspectives on language policy*. in *International Seminar on Constitutionalism and Diversity in Nepal*. 2007. Citeseer.
52. Galeoto, G., et al., *The Barthel Index: Italian Translation, Adaptation and Validation*. International Journal of Neurology and Neurotherapy, 2015. **2**: p. 2-7.
53. Kc, S., et al., *Nepali translation, cross-cultural adaptation and measurement properties of the Shoulder Pain and Disability Index (SPADI)*. Journal of Orthopaedic Surgery and Research, 2019. **14**(1): p. 284.
54. Nepal, G.M., M. Basaula, and S. Sharma, *Inter-rater reliability of Timed Up and Go test in older adults measured by physiotherapist and caregivers*. European Journal of Physiotherapy, 2019. **22**(6): p. 325-331.
55. Leung, S.O., C.C. Chan, and S. Shah, *Development of a Chinese version of the Modified Barthel Index-- validity and reliability*. Clin Rehabil, 2007. **21**(10): p. 912-22.
56. Hormozi, S., et al., *Iranian Version of Barthel Index: Validity and Reliability in Outpatients' Elderly*. Int J Prev Med, 2019. **10**: p. 130.
57. Minosso, J., et al., *Validation of the Barthel Index in elderly patients attended in outpatient clinics, in Brazil*. Acta Paulista de Enfermagem, 2010. **23**: p. 218-223.
58. Shah, S., *In praise of the biometric and psychometric qualities of the Barthel Index*. Physiotherapy, 1994. **80**: p. 769-771.
59. Hartigan, I., *A Comparative Review of the Katz ADL and the Barthel Index in Assessing the Activities of Daily Living of Older People*. International journal of older people nursing, 2007. **2**: p. 204-12.
60. Laake, K., et al., *The Barthel ADL index: factor structure depends upon the category of patient*. Age Ageing, 1995. **24**(5): p. 393-7.
61. Castiglia, S.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. **22**(4): p. 221-228.
62. Zhang, C., et al., *Psychometric properties of the Barthel Index for evaluating physical function among Chinese oldest-old*. JCSM Clinical Reports, 2022. **7**.
63. Küçükdeveci, A.A., et al., *Adaptation of the modified Barthel Index for use in physical medicine and rehabilitation in Turkey*. Scand J Rehabil Med, 2000. **32**(2): p. 87-92.
64. Tennant, A., J.M.L. Geddes, and M.A. Chamberlain, *The Barthel Index: an ordinal score or interval level measure?* Clinical Rehabilitation, 1996. **10**(4): p. 301-308.
65. DIN, P.C.B., *International classification of functioning, disability and health*. 2001.
66. Kline, P., *Handbook of Psychological Testing, second edition*. 2000.
67. Shah, S., F. Vanclay, and B. Cooper, *Improving the sensitivity of the Barthel Index for stroke rehabilitation*. J Clin Epidemiol, 1989. **42**(8): p. 703-9.

68. Hsueh, I.P., et al., *Comparison of the psychometric characteristics of the functional independence measure, 5 item Barthel index, and 10 item Barthel index in patients with stroke*. J Neurol Neurosurg Psychiatry, 2002. **73**(2): p. 188-90.
69. Quinn, T.J., P. Langhorne, and D.J. Stott, *Barthel index for stroke trials: development, properties, and application*. Stroke, 2011. **42**(4): p. 1146-51.
70. Terwee, C.B., Bot, S.D., de Boer, M.R., et al. (2007) Quality Criteria Were Proposed for Measurement Properties of Health Status Questionnaires. Journal of Clinical Epidemiology, 60, 34-42., *Quality Criteria Were Proposed for Measurement Properties of Health Status Questionnaires. Journal of Clinical Epidemiology*. . (2007. **60**: p. 34-42.
71. Nunnally, J.C.M.-H., *Psychometric theory*. 2nd Edition. 1978.
72. Post, M.W., et al., *[Dutch interview version of the Barthel Index evaluated in patients with spinal cord injuries]*. Ned Tijdschr Geneesk, 1995. **139**(27): p. 1376-80.
73. Cincura, C., et al., *Validation of the National Institutes of Health Stroke Scale, modified Rankin Scale and Barthel Index in Brazil: the role of cultural adaptation and structured interviewing*. Cerebrovasc Dis, 2009. **27**(2): p. 119-22.
74. Fricke, J. and C. Unsworth, *Inter-rater reliability of the original and modified Barthel Index, and a comparison with the Functional Independence Measure*. Australian Occupational Therapy Journal, 2010. **44**: p. 22-29.
75. Loewen, S.C. and B.A. Anderson, *Reliability of the Modified Motor Assessment Scale and the Barthel Index*. Phys Ther, 1988. **68**(7): p. 1077-81.
76. Galeoto, G., et al., *Evaluation of the psychometric properties of the Barthel Index in an Italian ischemic stroke population in the acute phase: a cross-sectional study*. Functional neurology, 2019. **34**: p. 29-34.
77. Duffy, L., et al., *Reliability (Inter-rater Agreement) of the Barthel Index for Assessment of Stroke Survivors Systematic Review and Meta-analysis*. Stroke; a journal of cerebral circulation, 2013. **44**.
78. Shrestha, M., et al., *Improving the Follow-up Rate for Pediatric Patients (0-16 years) of an Eye Hospital in Nepal: Protocol for a Public Health Intervention Study*. JMIR Res Protoc, 2021. **10**(10): p. e31578.
79. Kwakkel, G., et al., *Diagnostic accuracy of the Barthel Index for measuring activities of daily living outcome after ischemic hemispheric stroke: does early poststroke timing of assessment matter?* Stroke, 2011. **42**(2): p. 342-6.
80. *National Population and Housing Census 2021*. National statistics office, Government of Nepal: Kathmandu.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendix1.pdf](#)
- [Appendix2.pdf](#)
- [Appendix3.pdf](#)
- [BINPtableforclinicaluse.pdf](#)