

Reliability and Validity of the Nepali Wolf Motor Function Test following Cross-cultural Adaptation

Adhikari SP,^{1,2} Tretriluxana J,¹ Chaiyawat P¹

¹Faculty of Physical Therapy

Mahidol University, Thailand.

²Department of Physiotherapy

Dhulikhel Hospital, Kathmandu University Hospital

Dhulikhel, Kavre, Nepal.

Corresponding Author

Jarugool Tretriluxana

Deputy Dean for Academic and Research Affairs

Faculty of Physical Therapy, Mahidol University,
Thailand

Email: jarugool.tre@mahidol.ac.th

Citation

Adhikari SP, Tretriluxana J, Chaiyawat P. Reliability and validity of the Nepali wolf motor function test following cross-cultural adaptation. *Kathmandu Univ Med J* 2016;53(1):3-8.

ABSTRACT

Background

The Wolf motor function test is a standardized, reliable and valid performance-based measure, which evaluates upper-extremity function. Neither there is any evidence of neurological measure in Nepali nor has this tool been cross-culturally adapted to date.

Objective

To cross-culturally translate and evaluate reliability and validity of Nepali Wolf motor function test.

Method

Two forward-translators translated the original version into Nepali. After synthesizing, two back-translators translated it into English. The expert committee consolidated and derived a final Nepali version. Two assessors pretested on three participants to confirm that the original concept of the tool is preserved. The agreement and correlation between back-translators were evaluated. The inter- and intra-rater reliability and agreement of two physiotherapists on eight participants were demonstrated through intra-class-correlation-coefficient and weighted-kappa for time and functional ability respectively. Criterion validity was explored against Fugl-Meyer assessment scale.

Result

A cross-culturally adapted Nepali Wolf motor function test was pretested and ensured that the Nepali version was still retaining its equivalence (rho: 0.74 to 1.0 for time, 0.82 to 1.0 for functional ability). Good-to-excellent inter- and intra-rater reliability {intra-class-correlation-coefficient(2,1) and (3,1): 0.80 to 1.0 and 0.81 to 1.0 respectively} were demonstrated. Excellent agreement (kappa 0.90, $p < 0.00$) and good correlation (rho 0.57 to 1.0) between back-translators were found. An adequate criterion validity (rho -0.95 for time, 0.91 for functional ability) against Fugl-Meyer assessment scale has been demonstrated.

Conclusion

A cross-culturally adapted Nepali Wolf motor function test, preserving its original concept, was developed, and the reliability and validity of the tool on individuals with stroke has been demonstrated.

KEY WORDS

Reliability, stroke, translation, validity, wolf motor function test

INTRODUCTION

The Wolf Motor Function Test (WMFT) is a standardized, condition specific; performance-based upper extremity (UE) behavioral outcome measure. The modified WMFT consisting of seventeen sequential tasks; fifteen timed-items and two strength-items, evaluates UE performance through; a) timing, the speed at which functional task is completed, b) functional ability (FA), the movement quality and c) strength, the ability to lift against gravity.¹⁻⁴

This tool has been widely used as an outcome measure in participants with stroke, because; a) it is best suited for forced administered interventions b) it quantifies UE movement ability through performance time and functional tasks, c) its psychometric properties in varied level of stroke have been well established, (excellent inter- and intra-rater reliability; intra-class correlation coefficient 0.97-0.98, adequate concurrent, content, and construct validity, high internal consistency, specific minimal clinically important difference (time:1.5 to 2.0 seconds, FA: 0.2 to 0.4) and d) it consists of wide-ranged tasks in order of progressive complexity.²⁻⁹

WMFT has been translated to many languages such as Brazilian, French, Norwegian, Thai, and excellent psychometric properties are established.^{1,10-12} In spite of its wide use, appropriateness, feasibility, and well established psychometric properties, it has not yet been translated to Nepali language. There is no evidence of any neurological measure in Nepali culture and background to date. So, it is an urgent need for an appropriate outcome measure in Nepali context. Therefore, this study aimed to cross-culturally translate the WMFT into Nepali language and assess its reliability and validity in individuals with stroke.

METHODS

Translation with cross-cultural adaptation (First part of this study)

A written consent was received from the WMFT developer. The translation and cross cultural adaptation was done following the guidelines given by Beaton DE et al. because 1) this guideline is based on medical, sociological and psychological cross-cultural adaptation, 2) it has given emphasis to semantic, idiomatic, experimental and conceptual equivalence, 3) it is a refined guideline based on methodological experience in cross-cultural adaptation of generic and disease specific instruments and 4) it has been already adapted in translation of many tools including WMFT.^{1,13,14} The steps are briefly described in figure 1.

Stage I: Forward translation

Two native Nepali speakers, good in English (TOEFL and GRE score) were selected as forward translators (FT1 and FT2). One of the translators (FT1) was from health related

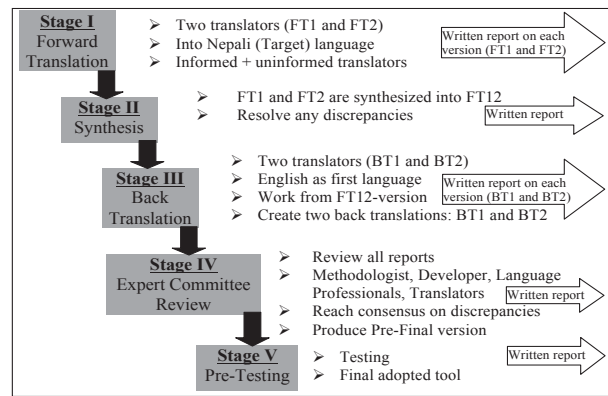


Figure 1. Steps of cross-cultural adaptation (Beaton DE et al. 2000, SPINE Volume 25)¹³ FT1: First forward translator, FT2: Second forward translator BT1: First backward translator, BT2: Second backward translator.

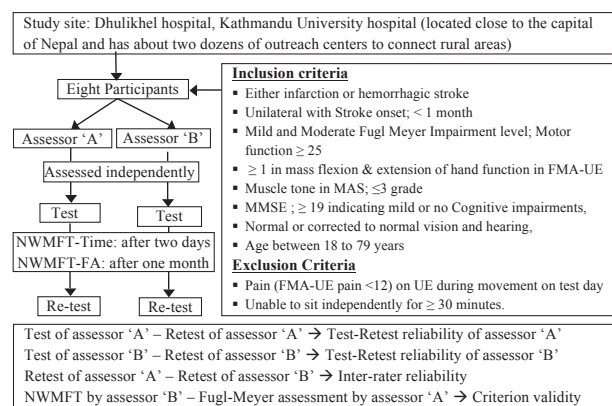


Figure 2. Procedural framework during reliability and validity testing.

field who was unaware of the original WMFT but familiar to the terminologies used in the tool whereas another (FT2) was professionally from literature background and was completely unaware of WMFT and its terminologies. They independently translated from the original language (English) to target language (Nepali). At times, when they were confused with the task/sentences, video was shown to them to clarify. There were some issues in finding exact Nepali words, and such issues were brought to the discussion and solved. A written report of the discussion has been prepared.

Stage II: Synthesis

Two forward translators and the researcher discussed together on both forward translations to produce one common forward translation (FT 12). Both translators did not find exact and suitable Nepali words for 'ready', 'set' 'go' and both agreed to replace them with 1, 2, 3. They also had some issues in finding Nepali word for 'basket', extremity', 'forearm' 'filming position', 'checkers', heel of hand, back of the body, and 'Dynamometer'. Those discrepancies were solved through discussion in number of rounds. It was emphasized that, the consensus between translators should be reached without compromising any of their feelings. Finally a common forward translation

(FT12) was developed. A detailed report of the process has been prepared.

Stage III: Back translation

Two back translators (BT1 and BT2); native English speakers capable in both spoken and written Nepali language were selected. Both of them had worked in Nepal for a number of years and had experience in translation. They were unaware of WMFT. The purpose of this back translation was for validity checking to make sure that the translated version is reflecting the same item content as the original version. Both of them faced some difficulties in understanding few typical Nepali words and their meaning had to be clarified through videos and pictures. At the end of this process, two back translations were produced. A written report of this whole process has been prepared.

Stage IV: Expert committee review

The composition of this expert committee comprises of methodologists, health professionals, language professionals and the translators.¹³ All the members were provided with both the forward as well as back translations together with the written report of each step. Some words/sentences were reanalyzed through number of rounds to resolve the discrepancies. Strong input was received from methodologists and language professionals. Once the committee was sure that the original concept of the tool was preserved, they consolidated and produced a semantic, idiomatic, experience and conceptual equivalent Nepali version of WMFT (NWMFT) through their critical decision.¹³ The written report of this step has also been prepared.

Stage V: Testing of the translated version (Pre-Testing)

Two physiotherapists (PT), one with and another without previous experience regarding administration of WMFT, assessed three patients, for both time and functional ability depending fully on NWMFT protocol. It was decided in advance to consider the lower level of the score in case any confusion on scoring.¹ The performance time and functional ability scores of two assessors, for each task, were compared and no significant difference ($p < 0.05$) was found. The Spearman's correlation between two assessors was found between 0.74 to 1.00 for NWMFT-time and 0.82 to 1.00 for NWMFT-FA respectively for different items. Perfect correlation and no difference in scoring between assessors proved that the translated version was clear in the target language and the concept of the original version was preserved.

The kappa as an agreement between two back translators has been analyzed. Despite some discrepancy in few terminologies/sentences as shown in table 1, an excellent agreement ($K = 0.90$, $p < 0.000$) was found. After clarifying those confusions through videos and pictures, a common consensus for all items was reached at the end without compromising anyone's feelings. Thus, a hundred percent agreement between the translators was found.

Table 1. Agreement between two back translators (BT1 and BT2)

Both BT1 & BT2 differ from OV (Original Version) on the following words/sentences	BT1 differs from OV	BT2 differs from OV	Kappa
Introducing test	Until the weight crosses the line	Keep shoulders level	0.90 ($p < 0.000$)
You can work up to 2 minutes			
The heel is beyond the front edge of the box			
Difficulty on finding Nepali word for; heel of hand, extremity, filming position, hand flat, checkers, bedside table, arm, forearm, elbow, wrist, back of body and handle	Dynamometer set on 2 nd setting		

BT1: First backward translator, BT2: Second backward translator, OV: Original version

Submission of the documentation to developers.

The detailed report and Nepali version of WMFT will be submitted to the original tool developer.

Testing reliability and validity of NWMFT (Second part of this study)

This study was a repeated measure design, which, was carried out on Nepalese individuals with stroke at varied levels of motor function and duration. Participants were selected by consecutive approach of convenience sampling technique. Ethical approval was received from the Institutional Review Committee of Kathmandu University School of Medical Sciences (IRC-KUSMS) (Approval Number 83/13). Informed written consent was also taken from each participant.

Statistical analysis:

The intra-class correlation coefficient {ICC(2,1) for inter-rater reliability and ICC(3,1) for test-retest reliability} was analyzed to evaluate the reliability for performance time and Spearman's rho was calculated to analyze correlation. Mann Whitney U test as well as kappa and descriptive statistics were analyzed using SPSS (version 19.00). Weighted kappa to analyze agreement for NWMFT-FA was calculated using STATA (version 12). The significant level was considered for $p < 0.05$.

RESULTS

The eight participants were heterogeneous with respect to age (mean 61.88 years), stroke type, territory involved, chronicity, muscle tone and level of motor as well as cognitive functions as shown in table 2. The agreement between two assessors and consistency between two tests of both assessors for NWMFT-FA score that were analyzed using weighted kappa are presented in table 3. The weighted kappa ranged from 0.52 to 0.92 between assessors, from 0.55 to 1.00 and 0.58 to 1.00 between test-retest for assessor 'A' and 'B' respectively. The result shows no difference ($p < 0.05$) between assessors and between tests as well.

Table 2. Clinical characteristics of the participants.

Subjects	Age (y)/gender	Stroke type/Territory involved	Onset (month)	MMSE score	FMA-UE score		MAS grade
					Motor	Total	
01	47/M	Hemorrhagic, MCA	3.5	28	48	92	1
02	51/M	Infarction, MCA	4	27	44	84	1+
03	54/M	Infarction, MCA	7	28	33	64	2
04	73/M	Infarction, MCA + ACA	4	22	38	74	1
05	61/F	Hemorrhagic, MCA	8	24	32	62	3
06	74/M	Infarction, MCA	9	24	30	58	2
07	57/F	Infarction, MCA	6	23	26	42	2
08	78/M	Infarction, MCA + PCA	10	20	25	38	3

MMSE: mini mental state examination, FUA UE- Fugl-Meyer assessment of upper extremity, MAS- modified Ashworth scale, y-year, M- male, F- female, MCA/ACA/PCA: middle/anterior/posterior cerebral artery.

Table 3. Weighted kappa (K) showing agreement for NWMFT-FA

WMFT Tasks	Agreement between assessors		Consistency between tests of assessor 'A'		Consistency between tests of assessor 'B'	
	K	p-Value	K	p-Value	K	p-Value
Forearm to table (Side)	0.85	0.001	0.85	0.001	0.85	0.001
Forearm to Box (Side)	0.79	0.001	0.79	0.001	0.64	0.000
Extend elbow (side)	0.70	0.002	0.55	0.006	0.58	0.011
Extend elbow (weight)	0.52	0.013	0.69	0.001	0.69	0.001
Hand to table (Front)	0.75	0.000	0.75	0.000	0.58	0.001
Hand to box (Front)	0.60	0.000	0.81	0.000	1.00	0.000
Reach and Retrieve	0.77	0.000	0.80	0.000	0.73	0.001
Lift can	0.81	0.000	1.00	0.000	0.90	0.000
Lift pencil	0.80	0.001	0.90	0.000	0.70	0.004
Lift paper clip	0.77	0.001	0.69	0.002	1.00	0.000
Stack checkers	0.92	0.000	0.92	0.000	0.85	0.000
Flip cards	0.73	0.004	0.58	0.005	0.81	0.000
Turn key in lock	0.83	0.000	0.73	0.000	0.85	0.000
Fold towel	0.82	0.000	0.73	0.001	0.92	0.000
Lift basket	0.87	0.000	0.87	0.000	0.69	0.003

NWMFT: Nepali wolf Motor function test, K: weighted kappa

The intra-class correlation coefficients (ICC) demonstrating inter- and intra- raters reliability of NWMFT-time are presented in table 4. The inter-rater reliability for NWMFT-time ranged from 0.80 to 1.00 whereas the test-retest reliability ranged from 0.82 to 1.00 for assessor 'A' and 0.81 to 1.00 for assessor 'B' respectively. The inter-rater as well as intra-rater reliability of both assessors for total score was found 1.00. The internal consistency indicated by Cronbach's alpha ranged from 0.89 to 1.00, 0.91 to 1.00 and 0.89 to 1.00 for between assessor and between tests of two assessor 'A' and 'B' respectively. Thus, the WMFT-time score of the translated version was consistent and found no difference ($p < 0.05$) between raters as well as between tests.

When the NWMFT-total time scored by assessor 'B' was compared with FMA-UE total scores, evaluated by assessor

Table 4. Inter and Intra (Test-Retest) - rater reliability of NWMFT-Performance time.

NWMFT	Inter-Rater reliability (ICC 2,1)	Test-Retest reliability (ICC 3,1) of assessor 'A'	Test-Retest reliability (ICC 3,1) of assessor 'B'
Forearm to table (Side)	0.81	0.94	0.81
Forearm to box (Side)	0.89	0.86	0.82
Extend elbow (Side)	0.86	0.83	0.81
Extend elbow (Weight)	0.90	0.95	0.86
Hand to table (Front)	0.84	0.90	0.82
Hand to box (Front)	0.84	0.82	0.87
Weight to box	0.98	0.97	0.98
Reach and Retrieve	0.80	0.85	0.91
Lift can	1.00	1.00	1.00
Lift pencil	1.00	1.00	1.00
Lift paper clip	0.99	0.99	0.99
Stack checkers	0.99	0.99	0.99
Flip cards	0.99	0.99	1.00
Grip strength	0.98	0.98	0.89
Turn key in lock	1.00	0.99	0.99
Fold towel	0.99	0.99	0.99
Lift basket	0.96	0.87	0.92
Total Score	1.00	1.00	1.00

ICC: Intra-class correlation coefficient

'A', a consistent and similar pattern was found. As shown in figure 3, higher the FMAUE-total score, shorter is the total NWMFT-time score and vice versa. Although there is a slight variation in scores in participants with less severity, the pattern is relatively same in all participants. The criterion validity of NWMFT-time and NWMFT-FA against FMAUE-total score have also been calculated using Spearman correlation coefficient (ρ) and found -0.95 and 0.91 respectively. The negative value indicates inverse relationship between the two. In addition, Spearman's correlation between two assessors for both NWMFT-time as well as NWMFT-FA has also been analyzed and found between 0.57 to 1.00 and 0.75 to 0.99 respectively.

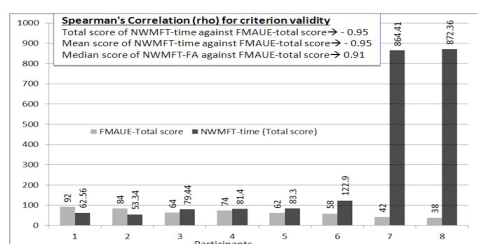


Figure 3. NWMFT-time and FA against FMA-UE total scores demonstrating criterion validity.

DISCUSSION

The WMFT has been cross-culturally translated to Nepali and its inter- and intra- rater reliability, consistency, correlation and the validity have been explored on heterogeneous participants. The cross-cultural adaptation was done fully depending on guidelines given by Beaton and colleagues.¹³ Detailed written report has been prepared on each step including discrepancies found and the ways they were solved. The consensus between translators was reached without compromising any of their feelings. Forward translators provided strong inputs at times of difficulty in finding exact and/or inappropriate Nepali words. The expert committee at the end produced an equivalent NWMFT through their critical decision ensuring that the original concept of the tool was retained. Pretesting of NWMFT was done to explore meaning of the items and responses.¹³ No significant difference was found on scoring between the previously trained and the newly trained assessors. An adequate correlation and excellent agreement between them were found on both domains, despite a few discrepancies, on which consensus was reached after clarification through pictures and videos. This has ensured that the NWMFT is still retaining its equivalence in applied situation.

The assessor 'A' (the main researcher of this study) who has been previously trained and the assessor 'B' (a registered, permanent, physiotherapist at government hospital having more than two years of experience in assessing and treating stroke patients) who has been recently trained for WMFT administration, tested the reliability and the validity of NWMFT on the same participants relying fully on its guidelines. Although selection of the assessors was made by convenience sampling, they represent physiotherapists of Nepal in general. The inter- and intra-rater reliability of NWMFT-time (direct assessment) was found good-to-excellent. These results are consistent with the findings of the original version, (intra- and inter- rater reliability between 0.97 to 0.99 and 0.92 to 0.99 respectively),^{3,4} as well the Brazilian version, (ICC for inter- and intra-rater reliability from 0.87 to 1.00).¹ The items measuring strength also showed consistent reliability similar to other tasks.¹⁻⁴ The weighted kappa for NWMFT-FA (on grading recorded video) showed good-to-excellent agreement between assessors and consistency between test and

retest of both assessors. Similar to reliability of NWMFT-time, the agreement for NWMFT-FA is also consistent with the findings of the original version,⁴ as well as Brazilian version ($k > 0.75$).¹ The weighted kappa of NWMFT-FA was found even higher than in a recent study by Duff SV et al. in which inter-rater reliability of WMFT focusing particularly to functional ability was tested.¹⁵ Similarly, this study showed findings consistent with the study by Moris et al. ($ICC \geq 0.90$),² and many other related studies and thus could establish adequate reliability and validity of NWMFT.^{6-9,14,15}

Good-to-excellent correlation found between assessors for both time and functional ability further supports the reliability and validity of NWMFT similar to a study.¹² Consistent to their study, we found relatively low rho for tasks 4, 5, and 8. The NWMFT-time and functional ability of two assessors were also compared using Mann Whitney U test which indicated no difference (p -value > 0.75) between assessors consistent with different literature findings.^{1,2,4} These findings and the literature evidence show that longer the time needed to complete a task, better is the reliability. Thus the items with increasing complexity increased reliability as seen even in other related studies.^{1,2,4} The intra-rater reliability was relatively higher than inter-rater reliability and the ICC of time was relatively higher than kappa for FA as evident in previous studies.²⁻⁴

Relying strictly on the adapted guidelines may give a good face validity of NWMFT.¹³ The excellent agreement between two back translators shows that NWMFT is valid enough and reflects the same item content as the original version, as evident in related studies.^{1,6-9,14,15} Since there is no similar neurological tool in place till the date in Nepali language, the criterion validity of NWMFT has been compared against English version of FMAUE, which is a standard, well-established performance tool, and has been used to compare criterion validity of WMFT in previous studies.^{4,9} The raw data of FMA-UE and NWMFT of both assessors showed similar pattern in all participants in spite of varied level of clinical features, which ensured criterion validity of NWMFT. The high correlation coefficient of both time and FA of NWMFT against FMAUE-motor score further supports its criterion validity consistency with a study by Whittall J and colleagues.⁹ In addition; the relationship between the scores of different tests further supports the criterion validity of the test as evident in the literature.⁴

Excellent reliability and agreement obtained for previously and newly trained assessors could be one of the strengths of this study. We could not evaluate other psychometric properties of NWMFT due to small sample size, which is a limitation and is therefore, recommended for future study.

CONCLUSION

A cross-culturally adapted NWMFT has been developed and an adequate reliability and validity on Nepalese individuals with stroke has been established. The ICC for

intra- and inter-rater reliability was found 0.80 to 1.00 for NWMFT-time. The weighted kappa for NWMFT-FA ranged from 0.52 to 0.92. Adequate criterion validity was found against FMA-UE (rho: -0.95 for time and 0.91 for functional ability). Thus, a Nepali version of WMFT retaining its equivalence with the original version has been produced through applied situation. It was tested and found reliable, valid, and feasible in Nepalese cultural background and context. It can be used as an evaluation tool in day-to-day clinical practice and research as well.

ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to forward translators (Dr. Hari Ram Adhikari, Ms. Rubee Dev), back translators (Dr. Darren Reed, and Dr. Stephen J. Knoble), Physiotherapist who was involved in testing (Mr. Madan Gautam), Physiotherapists of Dhulikhel Hospital and all participants without whom this study could not be possible.

REFERENCES

- Pereira ND, Michaelsen SM, Menezes IS, Ovando AC, Lima RC, Teixeira-Salmela LF. Reliability of the Brazilian version of the Wolf Motor Function Test in adults with hemiparesis. *Rev Bras Fisioter.* 2011 May-Jun;15(3):257-65. PubMed PMID: 21829991. Epub 2011/08/11. eng por.
- Morris DM, Uswatte G, Crago JE, Cook EW, 3rd, Taub E. The reliability of the wolf motor function test for assessing upper extremity function after stroke. *Arch Phys Med Rehabil.* 2001 Jun;82(6):750-5. PubMed PMID: 11387578. Epub 2001/06/02. eng.
- Wolf SL, Lecraw DE, Barton LA, Jann BB. Forced use of hemiplegic upper extremities to reverse the effect of learned nonuse among chronic stroke and head-injured patients. *Exp Neurol.* 1989 May;104(2):125-2. PubMed PMID: 2707361. Epub 1989/05/01. eng.
- Wolf SL, Catlin PA, Ellis M, Archer AL, Morgan B, Piacentino A. Assessing Wolf motor function test as outcome measure for research in patients after stroke. *Stroke.* 2001 Jul;32(7):1635-9. PubMed PMID: 11441212. Epub 2001/07/07. eng.
- Wolf SL, Thompson PA, Morris DM, Rose DK, Winstein CJ, Taub E, et al. The EXCITE trial: attributes of the Wolf Motor Function Test in patients with subacute stroke. *Neurorehabil Neural Repair.* 2005 Sep;19(3):194-205. PubMed PMID: 16093410. Epub 2005/08/12. eng.
- Edwards DF, Lang CE, Wagner JM, Birkenmeier R, Dromerick AW. An evaluation of the Wolf Motor Function Test in motor trials early after stroke. *Arch Phys Med Rehabil.* 2012 Apr;93(4):660-8. PubMed PMID: 22336104. Epub 2012/02/18. eng.
- Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. *J Hand Surg Am.* 1984 Mar;9(2):222-6. PubMed PMID: 6715829. Epub 1984/03/01. eng.
- Edward Taub DMMaJC. Wolf Motor Function Test (WMFT) Manual by UAB CI Therapy Research Group. 2011.
- Whitall J, Savin DN, Jr., Harris-Love M, Waller SM. Psychometric properties of a modified Wolf Motor Function test for people with mild and moderate upper-extremity hemiparesis. *Arch Phys Med Rehabil.* 2006 May;87(5):656-60
- Burge E, D. Kupper, M. Badan Ba, B. Leemann, A. Berchtold Qualities of a French version of the Wolf Motor Function Test; A multicenter study; *Annals of Physical and Rehabilitation Medicine* 56 (2013) 288-99.
- Vatshaug M. Inter-rater reliability of the Norwegian translation of the Wolf Motor Function Test; Institutt for Helse og Omsorgsfag Det Helsevitenskapelige fakultet Universitetet i Tromsø Oktober 2014.
- Tretriluxana J, Vachalathiti R, Pongvarin N, Emsakul J, Juntrkaew T, Srirugsa P, Thongkaew S. The reliability and validity of Wolf Motor Function Test (WMFT) for assessing paretic limb of individuals with acute stroke. *Journal of the Neurological Sciences.* 2009 Oct, Vol. 285, Supplement 1: S176-S177
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976).* 2000 Dec 15;25(24):3186-91. PubMed PMID: 11124735. Epub 2000/12/22. eng.
- Woodbury M, Velozo CA, Thompson PA, Light K, Uswatte G, Taub E, et al. Measurement structure of the Wolf Motor Function Test: implications for motor control theory. *Neurorehabil Neural Repair.* 2010 Nov-Dec; 24 (9):791-801.
- Duff S, He J, Nelsen M, Lane CJ, Rowe VT, Wolf SL, Dromeick AW, Winstein CJ. Interrater reliability of the Wolf Motor Function Test-Functional Ability scale: why it matters. *Neurorehabil Neural Repair.* 2014 Oct 16.