

University of Groningen

Reliability and validity of the Buruli ulcer functional limitation score questionnaire

Stienstra, Y.; Dijkstra, P.U.; van Wezel, M.J.; Beets, M.; Zijlstra, I.; Johnson, R.C.; Ampadu, E.O.; Gbovi, J.; Zinsou, C.; Etuaful, S.

Published in:
 American Journal of Tropical Medicine and Hygiene

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
 Publisher's PDF, also known as Version of record

Publication date:
 2005

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Stienstra, Y., Dijkstra, P. U., van Wezel, M. J., Beets, M., Zijlstra, I., Johnson, R. C., Ampadu, E. O., Gbovi, J., Zinsou, C., Etuaful, S., Klutse, E. Y., van der Graaf, W. T., van der Werf, T. S., & van Roest, M. H. (2005). Reliability and validity of the Buruli ulcer functional limitation score questionnaire. *American Journal of Tropical Medicine and Hygiene*, 72(4), 449-452.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

RELIABILITY AND VALIDITY OF THE BURULI ULCER FUNCTIONAL LIMITATION SCORE QUESTIONNAIRE

YMKJE STIENSTRA, PIETER U. DIJKSTRA, MARIEKE J. VAN WEZEL, MARGIJSKE H. G. VAN ROEST, MICHEL BEETS, IJSBRAND ZIJLSTRA, R. CHRISTIAN JOHNSON, EDWIN O. AMPADU, JULES GBOVI, CLAUDE ZINSOU, SAMUEL ETUAFUL, ERASMUS Y. KLUTSE, WINETTE T. A. VAN DER GRAAF, AND TJIP S. VAN DER WERF

Department of Internal Medicine and Department of Rehabilitation, Groningen University Hospital, Groningen, The Netherlands; Programme National de Lutte Contre l'Ulcère de Buruli, Ministère de la Santé Publique, Cotonou, Benin; Ministry of Health, National Buruli Ulcer Control Program, Accra, Ghana; Centre de Santé, Unité de Traitement des Ulcères de Buruli, Lalo, Benin; Centre Sanitaire et Nutritionnel Gbemontin, Zagnanado, Benin; Saint Martin's Catholic Hospital, Agroyesum, Ghana; Dunkwa Governmental Hospital, Dunkwa, Ghana

Abstract. The reliability and validity of the earlier developed Buruli ulcer functional limitation score (BUFLS) questionnaire was assessed. Of 638 former Buruli ulcer patients (of 678 individuals examined), sufficient items on daily activities (≥ 13 of the 19) were applicable to calculate a score. To determine the validity, the functional limitation scores of the 638 individuals were compared with the global impression of the limitations, range of motion (ROM), and the social impact (change of occupation or education) of Buruli ulcer. To determine inter-observer reliability, the functional limitation score was reassessed in 107 participants within one and three weeks after the first interview by another interviewer and interpreter. Both global impression and ROM correlated well with the functional limitation scores ($\rho = 0.66$ and $\rho = 0.61$). The inter-observer reliability of 107 participants as measured by an intra-class correlation coefficient of 0.86 was very good. The functional limitation scores measured in the second assessment were significantly higher than in the first assessment. This should be taken into account when the functional limitation score is used for the individual patient. The BUFLS can be used as for between group comparisons of endpoints in clinical trials and in the planning of resources.

INTRODUCTION

Buruli ulcer disease is the third most common mycobacterial disease after tuberculosis and leprosy.¹ It is caused by *Mycobacterium ulcerans*. According to the World Health Organization (WHO) clinical case definition, the pre-ulcerative stage includes nodules, plaques, or edema; in the ulcerative stage, skin ulcers with typically undermined edges can clinically be discriminated from other skin disorders. Later, a granulomatous healing response occurs, and fibrosis, scarring, calcification, and contractures with permanent disabilities may result.^{2–4} Occasionally, osteomyelitis complicates the course of illness. All these sequelae may lead to severe and permanent functional limitations. The current treatment recommendation implies extensive surgical treatment of Buruli ulcer lesions to ensure healing, yet extensive surgery may induce more extensive scarring and subsequent physical limitations.⁵

A reduction in the range of motion (ROM) was found in 58% of the former Buruli ulcer patients in Ghana when using a goniometer. Instead of measuring the physical limitations in ROM of joints, a simple and functional scoring system to assess nature and severity of the impairment on carrying out daily activities as a result of Buruli ulcer has been proposed.⁶ Based on frequency endorsement and applicability, a questionnaire with 19 items of daily activities was developed. The questionnaire has a good internal consistency ($\alpha = 0.89$).^{7,8} In this study, the validity and reliability of this questionnaire was analyzed.

METHODS

Population. In Ghana and in Benin, 678 former patients treated for Buruli ulcer were contacted based on records kept by the hospitals. These patients had finished treatment of Buruli ulcer in one of the five participating hospitals in Ghana

and Benin (Center de Santé, Unité de Traitement des Ulcères de Buruli, Lalo, Hospital Saint Camille, Davougou, and Center Sanitaire et Nutritionnel Gbemontin, Zagnanado in Benin, and Dunkwa Governmental Hospital, Dunkwa and Saint Martin's Catholic Hospital, Agroyesum in Ghana). Patients were enrolled in the study between January and June 2003. Included were patients more than three years old who had finished their treatment more than three months earlier. Excluded were patients spontaneously reporting to the study team who did not receive treatment in one of the hospitals participating in the study. All participants or their caretakers agreed to inclusion in the study. The study protocol was reviewed and approved by local hospitals and health authorities in Ghana and Benin.

Questionnaire. The questionnaire with 19 items of daily activities was used to assess the functional limitations of the former Buruli ulcer patients. These activities cover four dimensions: preparation of food/eating (four questions), clothing/personal care taking (three questions), working (five questions), and mobility (seven questions) (Appendix 1). The development of this questionnaire has been previously described in detail.⁷ Responses to each item are scored as 1 = "easily, on a normal level" if respondent could perform the activity without difficulties and on a level comparable to other community members of the same sex and age; 2 = "with difficulties" if respondent could perform the activity, but the level of performance is not the same as before Buruli ulcer started, the level is not comparable to other community members of the same sex and age, or the activity could be performed on the same level but only with difficulties; or 3 = "not possible at all" if a respondent could not perform this activity (without help of others) because of Buruli ulcer, both if physically impossible and if not possible because the respondent for example is avoiding the activity since he or she is afraid to damage the scar tissue.

If the item was not applicable for the respondent, e.g., the

person was too young or too old to perform that specific activity, the item was scored as "not applicable." For calculations of the individual functional limitation score, the number of answers "with difficulties" and "not possible at all" were divided by the number of activities applicable for that individual and multiplied by 100%. Thus, a higher score indicates more functional limitations. If more than 6 of the 19 items of the questionnaire were not applicable for a patient, no functional limitation score was calculated.

Validity. A global functional limitation score was assigned for each participant by means of the first visual impression the observer had of the patient before the questionnaire was used. Global impression on the functional limitations of the patients was categorized into no limitation, slightly limited, limited, severely limited, and very severely limited. The observer also assessed whether the participant had muscular atrophy at the affected body part. The occupation before and after the development of Buruli ulcer were asked for. If changes in occupation had occurred, the reason for the change was also asked for.

Range of motion. Both the affected joints and the joints not affected were measured in the participants. A total of 38 movements of the hand, wrist, elbow, shoulder, knee, and ankle were measured. Ranges of motion were recorded according to the SFTR (sagittal, frontal, and transverses rotation) method.⁹ For the assessment of the validity, the number of restricted motions was calculated. Restriction of motion was based on the values of the ROM according to Ellen and others.⁶

Reliability. The inter-observer reliability was assessed by performing a retest of the functional limitation score 1–3 weeks after the first assessment. The retest was performed by another interviewer and interpreter than in the initial test. For the analysis of the inter-observer reliability, the intraclass correlation coefficient (ICC) was used. The ICC expresses how well the two observers are likely to classify patients consistently relative to the other patients. The limits of agreement were used to provide an interval within which the differences between the repeated measurements are expected to lie.^{8,10–13}

Methodologic considerations on the calculation of the functional limitation score. In the earlier study on the functional limitation questionnaire,⁷ the functional limitation score was calculated by dividing the number of answers "with difficulties" and "not possible at all" by the number of activities applicable for that individual, and turned into a percentage. In this calculation, the functional limitations were dichotomized before calculating the percentage. The functional limitations can also be put on an ordinal scale, with 0 points if the activity was not limited, 1 point if the activity was "with difficulties," and 2 points if the activity is "not possible at all." The sum of the scores was then divided by the maximal score applicable for that patient. A higher score indicates more functional limitations with both calculations.

Data analyses. Internal consistency was analyzed by calculating Cronbach's alpha. An alpha value > 0.70 was considered sufficient. The correlation between the functional limitation score and the global impression, and the percentage of restricted motions of the limb affected was calculated by the Pearson correlation coefficient. In patients with one extremity affected, compensation mechanisms will take place by the other extremities. If patient have more than one extremity

affected, the possibility to compensate will change. Therefore, correlation with the percentage of restricted motions of the limb affected was assessed only for the patients with no more than one extremity affected to facilitate the interpretation. For the association between muscular atrophy and change of occupation because of Buruli ulcer with the functional limitation score, the Mann-Whitney U test was used.

For the inter-observer reliability, the ICC (one-way random) was used. An ICC > 0.75 was considered good. We planned to determine the limits of agreement and to plot the difference between the two sessions for each patient against the mean of each patient of the two sessions made. Data analyses were performed using SPSS version 10.0 (SPSS Inc., Chicago, IL).

RESULTS

Population. From January to June 2003, 678 former Buruli ulcer patients were included in the study. All patients had finished their treatment in one of the five participating hospitals. The characteristics of the patients are presented in Table 1.

Questionnaire. *General.* The questionnaire with 19 items had an internal consistency of $\alpha = 0.90$ for all 678 participants. When the participants with more than 6 of the 19 items not applicable were excluded, the internal consistency was $\alpha = 0.82$.

Validity. Global impression of the functional limitations of the patient and their functional limitation score (calculated ordinarily) showed a correlation ($\rho = 0.69$, $P < 0.001$) (Table 2). The average functional limitation score of patients with visible muscular atrophy was 26.4%, whereas the average functional limitation score of patients without visible muscular atrophy was 9.9% ($P < 0.001$, by Mann-Whitney U test). In the group of patients who had to change occupations or schools due to Buruli ulcer (i.e., financial problems, physical problems, or embarrassment), the mean functional limitation score was 25.6%, whereas the mean score in the group of patients who did not have to change their occupation or education due to Buruli ulcer was 9.0%. This difference was statistically significant ($P < 0.001$, by Mann-Whitney U test). The

TABLE 1
Baseline characteristics of 678 study participants*

Age in years at time of inclusion, mean	22
median (IQR)	15 (10–25)
Sex (male %)	49
Country, no. (%)	
Ghana	280 (41%)
Benin	398 (59%)
Length of time since end of treatment, mean in years (SD)	4.8 (2.6)
Amputation, no. (%)	26 (4)
Functional limitation scores (n = 638†)	
Dichotomized scoring mean (SD)	20 (24)
median	11
Ordinal scoring mean (SD)	14 (18)
median	5
Surgical treatment (%)	94
Antimicrobial treatment (%)‡	99

* IQR = interquartile range.

† 638 of the 678 participants who had ≤ 6 items of the questionnaire not applicable were included for the Buruli ulcer functional limitations score.

‡ Data were available for only 357 participants.

TABLE 2
Mean functional limitation scores and global impression

Global impression	Number (%) of patients	Mean functional limitation score
Not limited	413 (65)	5.9
Slightly limited	98 (15)	14.5
Limited	77 (12)	34.1
Severely limited	41 (6)	43.5
Very severely limited	8 (1)	48.4
Total	638 (100)	13.6

patients who did not have an amputation had a statistically lower functional limitation score than patients who did have an amputation (12.1% and 49.4%, respectively; $P < 0.001$, by Mann-Whitney U test).

Comparison with ROM. The percentage of restricted motions of the limb affected correlated moderately with the functional limitation score ($\rho = 0.61$, $P < 0.001$). Post-hoc analysis showed that the range of extension of the knee was limited in more than 40% of the knees not affected by Buruli ulcer disease had a limited ROM. Therefore, the ROM criteria used for the evaluation of the extension of the knee was adjusted (from -1° to -10°). After adjustment, only 5% of the unaffected knees had a limited ROM. Without this correction, the correlation between ROM and the functional limitation score was similar ($\rho = 0.56$, $P < 0.001$).

Inter-observer reliability. A total of 107 participants were reassessed. The mean \pm SD of the ordinal functional limitation score in the first and second sessions was 17.4 ± 20.4 and 21.4 ± 21.0 , respectively. The mean \pm SD difference was -4.0 ± 10.3 (95% confidence interval [CI] = 2.0 to 6.0). The ICC was 0.86 (95% CI = 0.80–0.90). Since the results of the second session were significantly higher than the results of the first session ($P < 0.001$, by Wilcoxon signed rank test), the limits of agreement could not be calculated.^{10,11}

Methodologic characteristics calculation of functional limitation score. The validity of the two different calculations of the functional limitation score was comparable (Table 3). The reliability showed a smaller mean \pm SD difference in favor of the ordinal calculation. The ICC for the two different ways of calculation of the functional limitation score was 0.88 (95% CI = 0.86–0.90).

DISCUSSION

To acquire a useful tool to assess nature and severity of the impairment caused by Buruli ulcer in carrying out daily ac-

tivities, the Buruli ulcer functional limitation score (BUFLS) was developed. In this study, the validity and reliability of the BUFLS were determined. The 19 items that were previously selected⁷ had a good internal consistency in this study with a large number of participants.

Validity of the questionnaire was good; the global impression of the functional limitations correlated well with the functional limitation score. The level of correlation between the global impression and the functional limitation score shows that the two do not measure the same construct. As expected, the functional limitation score was higher in participants with visible muscular atrophy and in participants who had an amputation. However, the functional limitation score is also associated with the social impact; participants who had to change their occupation due to Buruli ulcer had higher functional limitation scores. Initially, we planned to observe former patients during their daily activities to study the correlation between the observed score and the score as reported by the participants. Unfortunately, within the study period, it was not possible to observe sufficient daily activities of the participants.

The functional limitation score correlated moderately with the ROMs as measured. We believe that this correlation is moderate because the number of affected ROMs was included in the measurements, but the severity of restriction and the possible compensation mechanisms patients apply, were not included in the analysis. Furthermore, muscle strength, which also influences perceived limitations, was not measured. Normal ROMs have not been examined in African populations. Combined with the circumstances in which measurements have to be carried out, this may lead to some incorrect classifications of the ROMs measured. The coping mechanisms of the patient may also lead to a difference between ROMs measured and the self-reported functional limitation score.

The ICC of the inter-observer reliability was good. Scores of the retest were significantly higher than the scores of the first test. In other words, the participants reported to have more functional limitations when they were interviewed for the second time. Sociocultural factors may have contributed to this effect. This difference should be taken into account when using the BUFLS clinically. We believe the instrument should be used with caution when assessing functional limitations of individual patients.

The functional limitation score based on ordinal calculations has comparable validity and better reliability, probably since the severity of the limitations are better represented with the ordinal calculations. Therefore, this score should be used for further development of the scale.

In 1980, the WHO published an International Classification of Impairments, Disabilities and Handicaps Scoring System (ICIDH), which was succeeded by the International Classification of Functioning, Disability and Health.¹⁴ Leprosy research has based a scale to assess activities of daily living in persons affected by leprosy on the ICIDH-2. This scale performed well during validity and reliability testing, but is not applicable for Buruli ulcer patients.¹⁵ The Eye/Hand/Feet score, which is used as a scale to grade disability of leprosy, is also not applicable to Buruli ulcer patients because impairments of other parts of the body cannot be graded, and sensory impairment is not a common issue in Buruli ulcer.¹⁶

New treatment strategies and interventions should be de-

TABLE 3
Characteristics of the ordinal and dichotomized functional limitation scores*

	BUFLS dichotomized	BUFLS ordinal
Validity		
Correlation score/global impression	$\rho = 0.66$	$\rho = 0.69$
Correlation score/ROM	$\rho = 0.61$	$\rho = 0.55$
Reliability		
Intra-class correlation coefficient	0.82	0.86
Mean difference between test-retest (SD of difference)	5.8 (15.1)	4.0 (10.3)

* BUFLS = Buruli ulcer functional limitation score; ROM = range of motion.

veloped to prevent contractures and amputations or to rehabilitate the patients. After establishing the validity and reliability of the BUFLS, the responsiveness of the instrument to detect changes when patients are exposed to disease-modifying therapy should be assessed.¹⁷

Received June 1, 2004. Accepted for publication August 12, 2004.

Acknowledgments: We thank Ilona Hospers, Irene Wiersma, Valentin Gangbe, Adrien Hadjinde, Wilfried Houegnon, Gabriel A. Hagen, and Dennis Agyeman for their help with this study.

Financial support: This study was supported by the NWO (Dutch Organization for Scientific Research). Marieke J. van Wezel, Margijske H. G. van Roest, Michiel Beets, and IJsbrand Zijlstra received support funding from the Marco Polo fund, the Hendrik Muller's Vaderlandsch fund, the National Dermatology fund, and the Groningen University fund.

Authors' addresses: Ymkje Stienstra, Marieke J. van Wezel, Margijske H.G. van Roest, Michiel Beets, IJsbrand Zijlstra, Winette T. A. van der Graaf, and Tjip S. van der Werf, Department of Internal Medicine, Groningen University Hospital, PO Box 30.001, 9700 RB Groningen, The Netherlands, Fax: 31-50-361-3216, E-mails: y.stienstra@int.azg.nl and t.s.van.der.werf@int.azg.nl. Pieter U. Dijkstra, Department of Rehabilitation, Groningen University Hospital, PO Box 30.001, 9700 RB Groningen, The Netherlands. R. Christian Johnson, Programme National de Lutte Contre l'Ulcère de Buruli, Ministère de la Santé Publique, Cotonou, Benin. Edwin O. Ampadu, Ministry of Health, National Buruli Ulcer Control Program, Korle Bu, Accra, Ghana. Jules Gbovi, Centre de Santé, Unité de Traitement des Ulcères de Buruli, Lalo, Benin. Claude Zinsou, Centre Sanitaire et Nutritionnel Gbemontin, Zagnanado, Benin. Samuel Etuaful, Saint Martin's Catholic Hospital, Agroyesum, Ghana. Erasmus Y. Klutse, Dunkwa Governmental Hospital, Dunkwa, Ghana.

REFERENCES

- Asiedu K, Raviglione M, Scherpbier R, 1998. *International Conference on Buruli Ulcer Control and Research*. Yamoussoukro, Côte d'Ivoire, July 6–8, 1998. Geneva: World Health Organization.
- Stienstra Y, van der Graaf WTA, te Meerman GJ, The TH, de Leij LF, van der Werf TS, 2001. Susceptibility to development of *Mycobacterium ulcerans* disease: review of possible risk factors. *Trop Med Int Health* 6: 554–562.
- van der Werf TS, van der Graaf WTA, Tappero JW, Asiedu K, 1999. *Mycobacterium ulcerans* infection. *Lancet* 354: 1013–1018.
- Asiedu K, Scherpbier RW, Raviglione M, 2000. *Buruli Ulcer—Mycobacterium ulcerans Infection*. Geneva: World Health Organization.
- Teelken MA, Stienstra Y, Ellen DE, Quarshie E, Klutse E, van der Graaf WTA, van der Werf TS, 2003. Buruli ulcer: differences in treatment outcome between two centres in Ghana. *Acta Trop* 88: 51–56.
- Ellen DE, Stienstra Y, Teelken MA, Dijkstra PU, van der Graaf WTA, van der Werf TS, 2003. Assessment of functional limitations caused by *Mycobacterium ulcerans* infection: towards a Buruli ulcer functional limitation score. *Trop Med Int Health* 8: 90–96.
- Stienstra Y, Dijkstra PU, Guedenon A, Johnson RC, Ampadu EO, Mensah T, Klutse EY, Etuaful S, Deepak S, van der Graaf WTA, van der Werf TS, 2004. Development of a questionnaire assessing Buruli ulcer-induced functional limitation. *Am J Trop Med Hyg* 70: 318–322.
- Streiner DL, Norman GR, 1989. *Health Measurement Scales. A Practical Guide to Their Development and Use*. New York: Oxford University Press.
- Gerhardt JJ, Rondinelli RD, 2001. Goniometric techniques for range-of-motion assessment. *Phys Med Rehabil Clin N Am* 12: 507–527.
- Bland JM, Altman DG, 1999. Measuring agreement in method comparison studies. *Stat Methods Med Res* 8: 135–160.
- Bland JM, Altman DG, 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1: 307–310.
- Brouwer S, Kuijer W, Dijkstra P, Goeken L, Groothoff J, Geertzen J, 2004. Reliability and stability of the Roland Morris Disability Questionnaire: intra class correlation and limits of agreement. *Disabil Rehabil* 26: 162–165.
- Jensen MP, 2003. Questionnaire validation: a brief guide for readers of the research literature. *Clin J Pain* 19: 345–352.
- Ustun TB, Chatterji S, Bickenbach J, Kostanjsek N, Schneider M, 2003. The International Classification of Functioning, Disability and Health: a new tool for understanding disability and health. *Disabil Rehabil* 25: 565–571.
- van Brakel WH, Anderson AM, Worpel FC, Saiju R, Bk HB, Sherpa S, Sunwar SK, Gurung J, de Boer M, Scholten E, 1999. A scale to assess activities of daily living in persons affected by leprosy. *Lepr Rev* 70: 314–323.
- Brandsma JW, van Brakel WH, 2003. WHO disability grading: operational definitions. *Lepr Rev* 74: 366–373.
- Guyatt GH, Kirshner B, Jaeschke R, 1992. Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol* 45: 1341–1345.

APPENDIX 1

ITEMS ON THE BURULI ULCER FUNCTIONAL LIMITATION SCORE

Type of activity	Activity	Extremity involved
Preparation of food/eating	Fetching water from pump	Lower and upper
	Pound fufu (/manioc*)	Lower and upper
	Pouring water from a bottle into a glass	Upper
	Cutting vegetables with a knife	Upper
Clothing/personal care taking	Putting on T-shirt	Upper
	Wash yourself	Upper
	Cleaning yourself after using the toilet	Upper
	Using a cutlass	Lower and upper
Working	Heave loads on head	Lower and upper
	Carry harvest home	Lower and upper
	Opening bottle with screw top (/corked bottle*)	Upper
	Tie a knot	Upper
Mobility	Walking level ground	Lower
	Walking uphill	Lower
	Walking downhill	Lower
	Running	Lower
	Squatting	Lower
	Kneeling	Lower
	Standing up from floor	Lower and upper

* As asked in Benin.