Title page: The Norwegian General Motor Function assessment as an outcome measure for a frail elderly population: A validity study

Corresponding author:

Birgitta Langhammer, Professor PhD

Oslo and Akershus University College, Faculty of Health, Department of Physiotherapy

Posbox 4, St Olavs pl

0130 Oslo 4

Norway

Phone: + 47 22 45 25 10

Fax: + 47 22 45 25 05

e.mail: Birgitta.Langhammer@hioa.no

Co-author:

Birgitta Lindmark, Professor PhD

Uppsala University, Institution of Neurosciences, Physiotherapy Programme

Akademiska Sjukhuset ing 15

751 85 Uppsala

Sweden

+ 46 18 611 37 11

+ 46 18 50 19 89

e-mail: Birgitta.Lindmark@sjukgym.uu.se

Abstract

Aim: To establish validity of the Norwegian General Motor Function (NGMF) assessment scale.

Method: To establish construct and criteria validity Spearman's rank correlation coefficients (rs) was calculated for the NGMF and age, gender, medical conditions, the history of falls and to four functional tests. Content validity was evaluated by asking participating physiotherapists about the usefulness of the items in the scale. Absolute reliability was evaluated by establishing the Standard Error of Measurement (SEM) and the Minimal Detectable Change (MDC) at the 95% level of confidence for total scores of the NGMF subscales for dependence, pain and insecurity.

Results: Construct validity was established to medical status and medication with subscales dependence and insecurity but not to subscale pain. Criterion validity was established between the NGMF subscales dependence, pain and insecurity and the Barthel Index (BI), the Falls Efficacy Scale (FES) to subscales dependence and insecurity but not with pain, and the Timed Up-and-Go (TUG) test, to subscale insecurity. Neither the Chair-Stand Test nor registered falls were significantly associated to any of the subscales of NGMF. Content validity the NGMF was perceived relevant to the work in a geriatric setting and as a communication tool for a multidisciplinary team. MDC was calculated for dependence, 2.76, pain 4.9, and insecurity 6.1 respectively.

Conclusion: The construct, criteria and content validity of the NGMF was established.

Keywords: assessments; Norwegian General Motor Function Assessment Scale; older adults; validity

The Norwegian General Motor Function assessment as an outcome measure for a frail elderly population: A validity study

Introduction

An increasing population of older persons is maintaining their health and a good quality of life into their 80s and 90s and even beyond this age (1). Many examples of healthy elderly people who are independent, self-supporting and vitally active are presented in the media, but a divergence is also becoming visible. The division between people who are ageing "successfully" and people who are less successful seems to be related to both social circumstances and health issues (2-3). Frail elderly persons present three or more of characteristics: low physical activity, muscle weakness, slowed performance, fatigue, poor endurance and / or unintentional weight loss.

A major prerequisite for both health and quality of life is the physical ability to maintain independence in mobility and in the performance of the Activities of Daily Living (ADL) (4-5). It is well known that limiting activity is a risk factor for functional decline, morbidity and mortality among community-dwelling older adults (3-6). In particular, people with frail health are vulnerable to the risk of secondary conditions, as well as dysfunctions caused by a reduction in activity levels (4-8). Among older people, pain and the fear of falling are associated with limiting activity and a declining quality of life (QoL) (2).

The General Motor Function (GMF) assessment scale was constructed to meet the requirements of an instrument for screening and assessment of activity-related dependence, pain and insecurity among older people with frail health (9-10). The GMF has been translated into Norwegian, named the NGMF, and then tested for reliability but not validity (11).

The NGMF assessment – including its subjective aspects – is based on performance testing. It comprises the observation of dependency, with reports on pain and insecurity triggered by the execution of the tasks.

The intra- and inter reliability of the NGMF was analyzed, in a previous study, with a rank transformable method of ordered categorical assessments, developed by Svensson (19). Intra-reliability was found to be satisfactory with an augmented rank-order agreement coefficient (r_a) of 0.96 for Dependence, 0.96 for Insecurity and 0.99 for Pain. The augmented rank-order agreement coefficients for Inter-rater reliability were for Dependence $r_a = 0.97$; for Pain, $r_a = 0.99$; and for Insecurity, $r_a = 0.99$ (11).

Cronbach's alpha, commonly used as a measure of the internal consistency or reliability of a psychometric test score for a sample of patients, was also evaluated in this previous study (11, 25). The Cronbach's alpha coefficients were calculated for the different subscales of NGMF in the reliability study and were 0.68 for Dependency, 0.73 for Pain and 0.75 for Insecurity (11).

The results indicate that NGMF may be perceived as a relatively reliable instrument and considered useful, both in regard to the consistency of the scale and to intra- and inter-observer agreement, which also was in line with the original Swedish reliability test on a similar population of elderly people (9).

The aim of the present study was to evaluate the use of NGMF for construct, criterion and content, validity in frail geriatric community-based rehabilitation. The construct validity was evaluated for age, gender, medical conditions, and falls history. The criterion validity was evaluated by comparing the NGMF to four domains: the Barthel Index (BI) (12), the Timed-Up-and-Go (TUG) test (13), the Falls Efficacy Scale (FES) (14), and the Chair-Stand Test in 30 s (15-16). Finally content validity was evaluated for clinical usefulness and floor and ceiling effects.

Materials and Methods

Setting and subjects

Three physiotherapists, 30–60 years of age and with 5–40 years of experience, performed the tests. The inclusion criteria were frail, elderly, permanent or short-term residents in a nursing home or daycare users. The exclusion criteria were: persons with cognitive impairments not able to

understand instructions, or those with a terminal disease. Demographic data was obtained from the records held by the geriatric competence centre (Table 1).

The Regional Ethics Committee, 14 February 2011, no. 2010/3223, gave ethical approval. Information about the aim of the study was given both verbally and in writing to persons eligible for the study. Participation was voluntary and signed informed consent was obtained from each member of the sample group.

The General Motor Function Assessment

The NGMF is a performance-based test that assesses three factors affecting physical performance: dependence, pain and insecurity. A lower score indicates better performance, and 0 score indicates total independence / lack of pain or no insecurity.

The scale for mobility, 13 out of 21 items, is divided into three categories: 0=independent, 1= needs help from one person and 2= needs help from two people. By contrast, arm and hand items, eight out of 21, are divided into two categories: 0= independent and 1= needs help. Pain is reported for all the items. A score of 0 indicates that no pain was felt, whereas 1 indicates pain. Insecurity is reported with regard to the nine mobility items and two of the arm items (touch left/right toe), where 0 indicates no insecurity and a score of 1, insecurity (9-11).

Three separate total scores were calculated for dependence, ranging from 0 = independence, to 34 = total dependence; pain, ranging from 0 = no pain to 21 = serious pain, and insecurity, ranging from 0= no insecurity to 11= very insecure, respectively (11).

Previous tests of the NGMF concerning its clinical and psychometric properties for intra-reliability, inter-reliability and test-re-test reliability showed satisfactory overall results, while the field tests, interviews with physiotherapists in a geriatric setting, strengthened the evidence of its clinical feasibility by indicating clinical practicality, relevance and usefulness (11).

The Barthel Index

The Barthel Index (BI) is a test of performance of the activities of daily living, developed by Mahoney and Barthel (12) to measure functional independence in personal care and mobility. The scores reflect the amount of time and assistance a client requires. A score of 0 (complete dependence), 1, 2 or 3 is assigned to each level, with a possible total score of 20 (12, 17).

The Timed Up-and-Go

The Timed Up-and-Go (TUG) tests mobility, and is used in the clinic to evaluate dynamic balance, gait and transfers (13). The patient is asked to get up from a chair (46 cm high), with support for the arms, walk three meters, turn, go back and sit down. This is repeated twice and the best time is recorded. The physiotherapist monitors the time in seconds taken from the start, on getting up, to the end, when the patient is seated again (13).

Chair-Stand Test

The Chair-Stand Test (16) is a test of lower extremity strength. The patient is asked to stand up and sit down as many times as possible in 30 s. from a chair (46cm high) and the total number of stands is recorded. It is a part of the Senior Fitness Test (15-16), and normative values have been reported for healthy elderly persons (15-16).

Falls Efficacy Scale

The Falls Efficacy Scale (FES) is a self-evaluated test the participant's own confidence in their capacity to perform activities without fear of falling, developed by Tinetti (14). The participants rate 10 activities on a scale from 1 to 10, where 1 means total confidence in their own capacity to do the activity and 10 no confidence at all to perform the same activity. This means that higher scores represent less confidence in their performance of the activities without fear of falling, and low scores represent high confidence in their own capacity to perform the activities with no fear of falling. Cut off values >80 indicate increased risk of falling and > 70 indicate fear of falling (14).

Analysis of Validity

Construct validity

Construct validity refers to the extent to which a test measures the underlying theoretical constructs (18). To establish construct validity the association between NGMF and the theoretical constructs of a geriatric population (age, gender, medical conditions, and falls history) were analyzed.

Criterion validity

Criterion validity is established if the measure correlates well with other tests claiming to measure the same entity (18). Criterion validity, of the NGMF was compared to BI, TUG, Chair-Stand Test and FES by analyzing their correlation. All five outcome measures were assessed on the same patient on a single test occasion by a physiotherapist.

Content validity

Content validity, also known as logical validity, refers to whether or not the test covers a representative sample of items (18). In this study participating physiotherapists were asked about the usefulness of the items in the scale and of the scale in total. The results of this enquiry were presented in a previous study (11). Ceiling and floor effects were also examined in this earlier publication (11).

Statistical Methods

Spearman's rank correlation coefficient (r_s) was used to assess construct, criterion and content validity (25-26). Correlations between the BI, the Chair-Stand Test and the NGMF should result in a negative value, since a high score in the BI and the Chair-Stand Test indicate good performance, whereas a low score in the NGMF indicates a better performance (25-26). Correlations between the NGMF and the TUG, the FES, and registered falls, by contrast, should be positive since a lower score on all three indicate a better performance (25-26). Correlation coefficients were evaluated as <0.2 = low, 0.21-0.4= fairly low, 0.41-0.6= acceptable, 0.61-0.8 = good, and 0.81-1.0=very good (27). The Minimal Detectable Change (MDC) at the 95% level of confidence was evaluated for the NGMF

subscales for dependence, pain and insecurity, using the SEM values as follows: MDC= 1.96√2xSEM (19-24).

Results

A total of 30 elderly people, 20 women and 10 men, all in need of community-based services, were included in the study, (Fig 1). The majority were short-term patients in a nursing home, while a few others were living at home but receiving help from community-based services (Table 1). The mean age was 85 years (SD 7.4). All participants had multiple diagnoses were in need of multiple medications, were dependent on help with the ADL and had reduced walking capacity and mobility (Tables 1-2).

Construct validity

The NGMF scale and the subscale regarding dependence were significantly associated with medical status, and medication was associated with the subscale measuring insecurity (Table 3), indicating that the scale is associated with a geriatric population. However, neither age, nor gender, nor falls correlated to NGMF subscales for dependence, pain or insecurity (Table 3).

Criterion validity

The NGMF subscales for dependence, pain and insecurity correlated significantly with the BI, and the FES was significantly associated with the subscales for dependence and insecurity but not for pain (Table 4). The TUG test, however, was only significantly associated with the NGMF subscale for insecurity. Neither the Chair-Stand Test nor registered falls were significantly associated to any of the three NGMF subscales for dependence, pain or insecurity (Table 4).

Content validity

A focus group session with 12 participating physiotherapists, who had used NGMF in their clinical work in the Geriatric Centre, discussed the usefulness of the scale (11). They concluded that a benefit

of the NGMF was its focus on the simultaneous questions of pain and insecurity in combination with the tasks performed. The NGMF was perceived as a screening instrument, relevant to the work in a geriatric setting and as a communication tool for a multidisciplinary team (11).

Negative aspects were that the test took some time to perform: in general, between 15 and 45 minutes to conduct each test (Table 2).

Floor and ceiling effects

Reports on the floor and ceiling effects and the instrument's clinical usefulness have been made previously (11).The distribution of the NGMF scores showed that the total score was reached in approximately 50 % of the 30 subjects, which means that the majority of the participants could perform the items independently, indicating low sensitivity for the further improvement of performance. There were no floor effects in the three subscales, indicating that the subscales were able to discriminate between individuals with functional limitations and a disability (11).

Minimal Detectable Change

A Minimal Detectable Change (MDC) score was calculated for NGMF dependence, pain and insecurity. This rendered MDC for the subscales as follows: dependence, 2.8 points, pain 4.9 points, and insecurity 6.1 points.

Discussion

The main aim of this study was to establish the construct, criterion and content validity of the NGMF. In a previous study the internal consistency of NGMF and intra- inter reliability has been presented (11). Construct validity was correlated to medical status, criterion validity was shown by an acceptable association between the BI, the TUG and the FES and the subscales dependence, insecurity and pain, and content validity was appraised by the participating therapists and was considered adequate, indicating that the scale was valid for its purpose.

Construct

The subscale dependence was associated with medical status, geriatric diagnosis and medication, indicating that the construct of a geriatric population was established (Table 3). However, neither age, nor gender, nor falls correlated with any of the NGMF subscales for dependence, pain or insecurity. This might partly be explained by the small sample and that their age and gender were similar (Tables 1-2). Also the fact that participants were recruited from a geriatric rehabilitation setting, where patients were under supervision and the likelihood of falls was less than it might be in a home setting, may contribute to the low scores reported on the FES (Table 1-2).

Criterion validity

BI

The association to the performance of ADL that had previously been established in the Swedish version (9) was also established for the NGMF. The NGMF showed an acceptable association to the ADL measured with the help of the BI for dependence and pain and a fair association to total scores for insecurity (Table 4), which is in line with earlier studies (9-10). However, the association was lower (r = -0.58), than in the study made by Åberg et al. (r = 0.8) (28).This lower association could be explained by the fact that a different ADL outcome, the Katz Index of the ADL, was used by Åberg et al. (28-30). It may be that the Katz Index of the ADL shows a better fit because of the different selection of items than in the BI (28-30). However, this is pure speculation.

TUG

There was no association between NGMF for dependence or for pain and the TUG (Table 4), despite previous studies that had established such an association (10). One explanation may be that the sample consisted of relatively frail elderly people with poor walking capacity and problems with transfers (Tables 2). The moderate but significant correlation with NGMF insecurity may mirror the patients' worries about poor performances (Table 4).

The Chair-Stand Test

The Chair-Stand Test in 30 s. is a test of lower extremity strength, an important body function for performance of the ADL and transfer. The rationale is that lower leg strength may be directly

associated to the NGMF subscale for dependence, but this association was not established (Table 4). An explanation may be that the participants in this study were frail and weak in their lower extremities overall. Furthermore, the reported scores are considerably lower than normative age related values for healthy elderly Norwegians, where the Chair-Stand Test has been measured with a mean of 13.1 among healthy elderly people aged 80-89 years versus the mean of 2.6 in this study (15). One might speculate whether the group was too homogeneous, with respect to frailty (Table 2).

FES

The Falls Efficacy Scale was significantly associated with dependence (fair) and insecurity (good), which establishes NGMF as an important tool to predict falls for frail elderly people (Table 4). If cut points in the NGMF could be established it would be useful for clinical practice. Registered falls, by contrast, were not associated to NGMF. However, there were several difficulties related to the calculation and registration of falls. The falls were reported retrospectively during a period with a mean of 25 days. It was difficult for the therapists to get an accurate number since the reports came from the participant's own registration. Nevertheless, the few falls might suggest that the majority of the participants felt safe in the community health care surroundings, which was supported by the relatively low scores of the FES, indicating little or no fear of falling (Table 1).

Content

The participating physiotherapists considered the NGMF to be relevant in its function for geriatric multidisciplinary rehabilitation, especially the combined reports of pain and insecurity along with dependence scoring. This is in line with other reports (9-10). However the ceiling effect of the NGMF was considered a limiting factor in its use for geriatric rehabilitation of more "able" persons. However, the participants in this study were frail (Tables 1-2) and the NGMF can be perceived as a scale for frail elderly persons, in this respect.

Floor and ceiling effect

About half of the cases received a maximum total score, thus indicating a ceiling effect that had also been shown in the Swedish sample (9). This ceiling effect may limit the clinical usefulness of NGMF, suggesting that it should only be used for the frailest persons if the aim is to show improvement. However, for clinical use in which the aim is to monitor the maintenance of performance in the elderly in order to prevent deterioration, and as a tool for communication in multidisciplinary work in nursing homes and community care, the NGMF can be perceived as a useful and helpful instrument.

Minimal Detectable Change

The MDC for dependence was established at 2.8 points, for pain 4.9 points and insecurity 6.1 points respectively. The MDC means the minimum amount of change in a patient's score to ensure that the change is not the result of measurement error (23). These MDCs may be compared to a clinical study (29), reporting GMF change scores for dependence 3.1, pain 2.8 and insecurity 2.4 respectively. Unlike the MDCs reported in our study, these results might be borderline or too small to be considered real change scores (22). However, the calculation for the MDC is related to the mean, not the median, as reported in the intervention study (29). Our results though indicate limited discriminant ability of NGMF.

The Minimally Clinical Important Change (MCIC) it yet to be established and, for that to be accomplished, further and larger studies are warranted.

The results of our study must be considered with caution since a relatively small number of elderly people were included (n=30) which limits the generalizability of the results.

The NGMF was shown in this study to be a valid instrument, ranking from fair to good, for use with frail elderly persons in clinical practice with MDC scores for dependence (2.8), pain (4.9) and insecurity (6.1). NGMF subscales for dependence, pain and insecurity were significantly associated with performance in the activities of daily living measured with BI. Falls efficacy measured with the FES was associated with the subscales for dependence and insecurity, and the TUG with insecurity.

The clinical usefulness of the test was stressed in relation to its focus, the frail elderly, as was its value in communications between the members of the multidisciplinary team.

Acknowledgements

We would like to thank A-C Åberg, Associate professor and creator of GMF, and the Geriatric Competence Centre in Drammen, Norway for their valuable help and cooperation in this study. We would also like to thank, Pearson Assessment Sweden, the copyright holder of GMF, for allowing us to publish the results.

Declaration of interest

The authors report no conflict of interest.

References

1. Behm L, Wilhelmson K, Falk K, Eklund K, Zidén L, Dahlin-Ivanoff S. Positive health outcomes following health-promoting and disease-preventive interventions for independent very old persons: long-term results of the three-armed RCT Elderly Persons in the Risk Zone. Arch Gerontol Geriatr 2014; 58 (3):376-83.

2. Kempen GI, van Haastregt JC, McKee KJ, Delbaere K, Zijlstra GA. Socio-demographic, health-related and psychosocial correlates of fear of falling and avoidance of activity in community-living older persons who avoid activity due to fear of falling. BMC Public Health 2009; 02(.9): 170.

3. Doyle YG, Mc Kee M, Sherriff M. A model of successful ageing in British populations. Eur J Public Health 2012;. 22 (1):71-6.

4. Vermeulen J, Neyens JC, van Rossum E, Spreeuwenberg MD, de Witte LP. Predicting ADL disability in community-dwelling elderly people using physical frailty indicators: A systematic review. BMC Geriatric 2011; 11(33): 1-11.

5. Mercier L, Audet T, Hébert R, Rochette A, Dubois MF. Impact of motor, cognitive, and perceptual disorders on ability to perform activities of daily living after stroke. Stroke 2001; 32 (11): 2602-8.

6. Ueshima K, Ishikawa-Takata K, Yorifuji T, Suzuki E, Kashima S, Takao S et al. Physical activity and mortality risk in the Japanese elderly: A cohort study. Am J Prev Med 2010; 38 (4): 410-8.

7. Fletcher PC, Guthrie DM, Berg K, Hirdes JP. Risk factors for restriction in activity associated with fear of falling among seniors within the community. J Patient Saf 2010; 6 (3): 187-91.

8. Berlin Hallrup L, Albertsson D, Bengtsson Tops A, Dahlberg K, Grahn B. Elderly women's experiences of living with fall risk in a fragile body: A reflective life world approach. Health Soc Care Community 2009; 17 (4): 379-87.

9. Aberg AC, Lindmark B, Lithell H. Development and reliability of the General Motor Function Assessment Scale (GMF) - a performance-based measure of function-related dependence, pain and insecurity. Disabil Rehabil 2003; 25 (9): 462-72.

10. Gustafsson U, Grahn B. Validation of the General Motor Function Assessment Scale - an instrument for the elderly. Disabil Rehabil 2008; 30 (16):1177-84.

11. Langhammer B, Lindmark B. General motor function assessment scale - reliability of a Norwegian version. Disabil Rehabil 2014:36(20):1704-12.

12. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965; 14: 61-5.

13. Podsiadlo D, Richardson S. The Timed Up and Go – a test of basic functional mobility for frail elderly persons. J Am geriatric Soc 1991; 24: 398-401.

14. Tinetti ME; Richman D; Powell L. Falls efficacy as a measure of fear of falling. J Gerontol 1990; 45 (6): 239-243

15. Langhammer B, Stanghelle JK. Functional fitness in elderly Norwegians measured with the Senior Fitness Test. Adv Physiother2011; 13(4):137-144.

16. Rikli RE, Jones CJ. Functional fitness normative scores for community-residing older adults, ages 60 – 94. J Aging Phys Act 1999; 7:162 – 81.

17. Hsieh YW, Wang CH, Wu Sc, Chen PC, Sheu CF, Hsieh CL. Establishing the minimal clinically important difference of the Barthel Index in stroke patients. Neurorehabil Neural Repair. 2007;21:233–238.

18. French S, Reynolds F, Swain J. Practical research. A guide for therapists.2nd ed. Oxford, Butterworth and Heinemann, 2002.

19. Mao HF, Hsueh IP, Tang PF, Scheu CF, Hsieh CL. Analysis and comparison of the psychometric properties of three balance measures for stroke patients. Stroke 2002; 33:1022-7.

20. Holmes WC, Shea JA. Performance of a new, HIV/AIDS-targeted quality of life (HAT-QoL) instrument in asymptomatic seropositive individuals. Qual Life Res 1997; 6: 561-71.

21. Atkinson G, Neville AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217-38.

22. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. Phys.Ther 2006;86:735-43.

23. Wagner JM, Rhodes JA, Patten C. Reproducibility and minimal detectable change of a three dimensional kinematic analysis of reaching tasks in people with hemiparesis after stroke. Phys Ther 2008;88:652-63.

24. Schwenk M, Gogulla S, Englert S, Czempik A, Hauer K. Test-retest reliability and minimal detectable change of repeated sit-to-stand analysis using one body fixed sensor in geriatric patients. Physiol Meas 2012; 33 (11):1931-46.

25. Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1991.

26. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1: 307-310.

27. Cohen, J. Statistical power analysis for the behavioral sciences (2nd ed.). New Jersey: Lawrence Erlbaum 1988.

28. Åberg AC, Lindmark B, Lithell H. Evaluation and application of the General Motor Function assessment scale in geriatric rehabilitation. Disabil Rehabil 2003;25:360-68.

29. Åberg AC. Gender comparisons of function-related dependence pain and insecurity in geriatric rehabilitation. J Rehabil Med 2006; 38: 73-79.

30. Hartigan I. A comparative review of the Katz ADL and the Barthel Index in assessing the activities of daily living of older people. Int J Older People Nurs 2007; 2 (3): 204-12.

Figure legend:

Figure 1 Flow chart of the validity testing procedure

Table 1 - Description of the participating patients

	Subjects (n=30)
Age (mean; SD)	84.9 (7.4)
Female / male	86.0(5.9) / 82.3 (9.9)
Gender female /male	20 /10
Diagnosis (n):	
Hip fracture	3
Cerebral vascular accident	6
Cancer	2
Osteoporosis	1
Obstructive lung disease	2
Heart conditions	1
Reduced capacity	2
Other	13
Medications (n):	
0	3
1-5	13
6-10	9
11-12	5

Table 2 Age, diagnosis (Diag), Barthel Index (BI), Timed-Up-and-Go (TUG), Chair-Stand Test (Ch-St-t), Falls Efficacy Scale (FES), and Time for test procedures (min) were registered in mean (m) and standard deviation (SD). Medication (Med), diagnosis (Diag) and number of falls are presented in range of incidents. Results are presented for each of the testers and in total.

Tester	**Age**	**Med**	**Diag**	**Diag**	**BI**	**TUG**	**Ch-st-t**	**FES**	**Falls**	**Time**
1 **(n=10)**	86.2 (5.6)	0-12	7-11	9.2 (2.1)	17.3 (1.3)	33.8 (27.1)	2.3 (3.6)	12.1 (4.5)	0-2	23.8 (9.5)
2 **(n=10)**	83.6 (6.4)	0-10	7-11	9.0 (2.8)	15.3 (3.4)	31.3 (20.1)	3.2 (4.7)	13.6 (6.1)	0-2	20.8 (2.3)
3 **(n=10)**	84.1 (9.9)	3-12	2-8	4.0 (2.1)	16.1 (3.2)	34.0 (20.4)	2.6 (2.9)	9.4 (2.8)	0	35.0 (31.0)
All (n=30)	84.9 (7.4)	0-12	2-11	6.6 (3.4)	15.7 (3.1)	33.0 (19.8)	2.6 (3.5)	11.7 (4.8)	0-2	26.5 (19.2)

Table 3 Construct validity: Correlation coefficients between NGMF subscales for dependence, pain and insecurity and age, gender, geriatric diagnosis, falls history and medication, p<0.05.

	NGMF dependence	NGMF Pain	NGMF Insecurity
Age	0.1	0.1	0.1
Gender	- 0.2	- 0.1	- 0.1
Geriatric diagnosis	***0.7****[1]	- 0.2	- 0.3
Falls history	0.2	0.1	0.2
Medication	0.1	- 0.1	***0.4****

[1] * = p<0.05

Table 4 Correlation between functional outcomes, total score for BI, TUG, Chair Stand Test, Falls Efficacy Scale and reported falls and NGMF with $p<0.05$

	NGMF Dependence	**p-value**	**NGMF Pain**	**p-value**	**NGMF Insecurity**	**p-value**
Barthel Index tot	***-0.58***	***0.001****	***-0.48***	***0.009****	***-0.40***	***0.03****
TUG	0.2	0.3	0.3	0.12	***0.44***	***0.02****
Chair-Stand-Test	-0.08	0.72	-0.25	0.24	-0.25	0.24
FES tot	***0.38***	***0.04****	0.29	0.12	***0.64***	***0.001****
Falls	-0.12	0.55	-0.001	0.49	0.05	0.8

Figure 1 - Flow chart of the validity testing procedure