J Rehabil Med 2003; 35: 141–144

Taylor & Francis healthsciences

IS THE RIVERMEAD MOBILITY INDEX A SUITABLE OUTCOME MEASURE IN LOWER LIMB AMPUTEES?—A PSYCHOMETRIC VALIDATION STUDY

Franco Franchignoni,[1] Stefano Brunelli,[2] Duccio Orlandini,[3] Giorgio Ferriero[1] and Marco Traballesi[2]

From the [1]Unit of Occupational Rehabilitation and Ergonomics, Rehabilitation Institute of Veruno, "Salvatore Maugeri" Foundation, Clinica del Lavoro e della Riabilitazione, IRCCS, Veruno, [2]Fondazione S. Lucia, IRCCS, Rome and [3]Centro Protesi INAIL, Vigorso di Budrio, Italy

Objective: **To examine the internal consistency, validity, responsiveness and test scalability of the Rivermead Mobility Index.**

Design: **Methodological research (consecutive sampling, prospective longitudinal study).**

Patients: **140 unilateral lower limb amputees (79 above-knee and 61 below-knee).**

Methods: **The Rivermead Mobility Index was administered to all patients at the beginning (T_0) and at the end (T_2) of the prosthetic training. In 70 of the patients, the Functional Independence Measure and a timed walking test were also carried out.**

Results: **The Cronbach's alpha of the Rivermead Mobility Index was 0.85 and the item-to-total correlation coefficients r_{pb} ranged from 0.33 to 0.74 ($p < 0.0001$), for the items considered, at T_0; 4 correlations were not calculated due to the extremely low variability of some item responses (mode $\geq$98%). The correlation (r_s) of Rivermead Mobility Index score with the motor subscale of the Functional Independence Measure was 0.83 at T_0 and 0.69 at T_2 ($p < 0.0001$, for both) and that with timed walking test -0.70 ($p < 0.0001$) at T_2. The effect size was 1.35. The scalability coefficients were below the limits of acceptability.**

Conclusion: **When applied in lower limb amputees, the Rivermead Mobility Index is an ordinal measure with adequate levels of a series of psychometric properties, which seems more useful for epidemiological studies than for clinical decision-making in single patients. Further steps should be considered to improve its item selection, response format and scaling properties.**

Key words: amputees, leg prosthesis, outcome assessment, psychometrics

J Rehabil Med 2003; 35: 141–144

Correspondence address: Franco Franchignoni, Fondazione Salvatore Maugeri, Clinica del Lavoro e della Riabilitazione, Via Revislate 13, IT-28010 Veruno (NO), Italy. E-mail: ffranchignoni@fsm.it

Submitted August 14, 2002; accepted December 2, 2002

INTRODUCTION

Mobility is one of the most important outcomes in the rehabilitation of lower limb amputees (LLA), but there is no consensus about the best tool for measuring this variable (1). Among the many rating scales and questionnaires available to assess mobility, the Rivermead Mobility Index (RMI) (2), a cumulative index containing 15 dichotomous items, is one of the most widely adopted (3). The instrument has been validated in a variety of neurological diseases (4–8), and has recently also been used in LLA to assess mobility changes related to rehabilitation treatment (9, 10). There is, however, requirement for evidence of acceptable levels of a series of psychometric and practical properties within LLA before this index can be recommended as a suitable outcome measure for research studies or clinical applications in this particular context (11).

For this reason, the aim of the study was to examine the internal consistency, validity, responsiveness and test scalability of the RMI, as well as its appropriateness and precision, when applied in LLA, in order to assess the advantages and disadvantages of this specific use of the instrument.

MATERIAL AND METHODS

Subjects

One hundred and forty patients (102 males and 38 females; mean age 57 years, SD 18 years) participated in the study. They were admitted from 1st September 1999 to 31 May 2001 to 3 free-standing Rehabilitation Centres for rehabilitation and prosthetic training following a recent unilateral lower limb amputation. Patients affected by dementia, severe neurological, cardiac and respiratory diseases, or serious residual limb deformations were excluded. Amputation aetiologies were: peripheral vascular diseases (52.9%), trauma (32.1%), tumour (12.1%) and infective diseases (2.9%). Level of amputation was above-knee in 81 cases (58%) and below-knee in 59 cases (42%).

The study was approved by the local Ethics Committee. All patients gave their informed consent before entering the study.

Instruments and procedure

The Rivermead Mobility Index (RMI) was administered to all patients at the beginning (T_0) and at the end (T_2) of the prosthetic training. Furthermore, in 70 of the patients, randomly selected, 2 other measurements were carried out: the Functional Independence Measure (FIM) at T_0 and T_2 (simultaneously with the RMI), and a timed walking test (TWT), performed as soon as the patient began to walk outside of the parallel bars (T_1) and at T_2.

The RMI (2) includes 15 mobility items: 14 are questions put directly to the patient, while one (standing unsupported) relates to an observed performance. Dichotomous "yes/no" answers are scored 1/0 and then summated; hence the cumulative score may range from 0 to 15, with a higher score indicating better patient mobility.

The FIM (12) is an ordinal scale composed of 18 items with 7 levels of response ranging from 1 (total dependence) to 7 (total independence) designed to determine the level of disability of patients, as reflected by their need for assistance in daily activities. The scale can be subdivided into a 13-item motor subscale (motFIM) and a 5-item cognitive subscale (cognFIM). The ranges of scoring are motFIM: 13-91; cognFIM: 5-35. Lower scores denote a poorer performance.

 ISSN 1650–1977

The TWT (13) assesses the time needed by the patient to walk 10 metres on the level in a straight line, using their own walking aid, if any, and at their own preferred speed. The time taken from the "go" order to complete the 10-metre walk is measured (in seconds) with a stopwatch. The mean of 2 consecutive trials was considered in our study.

Statistics

The median was used as a measure of the central tendency of the distribution of the RMI scores, and the interquartile range (the range between the 25th and 75th percentiles) as a measure of their variability. The reliability (in terms of internal consistency) of the RMI was assessed at T_0 by means of: (*i*) the Cronbach coefficient alpha and (*ii*) the item-to-total correlation examining how well each of the RMI items was correlated (point biserial correlation coefficients, r_{pb}) with the total score, omitting that item from the total (14, pp. 575–577).

Convergent construct validity of the RMI was analysed through correlation (Spearman rank correlation coefficient, r_s) with scores of motFIM and TWT, simultaneously collected, hypothesizing a moderate/good (>0.50) to good/excellent (>0.75) correlation (14, p. 494) with both, based on the assumptions that the FIM describes the patient's independence of external help in basic activities of daily life and that the TWT explores the speed of walking a short distance, 2 variables related to mobility.

The responsiveness of the RMI was evaluated calculating: (*i*) the change scores, reported in terms of median (and interquartile range) of the individual differences between the T_2 and T_0 scores. In accordance with Hsieh et al. (7), a progress of 3 or more points on the RMI was considered as a clinically significant improvement; (*ii*) the correlation (r_s) of the RMI change scores with the correspondent changes in motFIM (11, p. 29) in 70 of the patients (see "Instruments and procedure"), during the observation time window; (*iii*) the effect size, defined as mean change score (T_2 minus T_0) divided by the standard deviation of the T_0 scores (15) (according to Cohen's suggestion, scores below 0.4 were considered small, a score of 0.5 moderate, and scores 0.8 or more large (14, p.105)), in the entire sample.

The fit of the RMI with the Guttman scaling (requiring that items represent a stable hierarchy of increasing amounts of an underlying trait) was calculated through:

(i) the coefficient of reproducibility (CR), with the Goodenough-Edwards technique, expressed as follows:

$$CR = 1 - \text{Scaling Errors (SE)}/\text{Total Responses (TR)}$$

where TR = number of items × number of subjects, and the SE (number of deviations of the observed response pattern from the pattern required by the Guttman's model) are calculated rearranging item columns to minimize them;

(ii) Menzel's coefficient of scalability (CS), with the formula:

$$CS = 1 - \text{Scaling Errors (SE)}/\text{Marginal Errors (ME)}$$

where ME are the sum of all non modal frequencies and CS represents a measure of the scale's ability to predict item responses in comparison to predictions based on marginal frequencies.

Conventionally, acceptable values for CR should be $\geq$0.90 and for CS $\geq$0.65 (16).

RESULTS

The median (and interquartile range) of the RMI scores in the 140 patients was 7 (5–9) at T_0, and 11 (10–13) at T_2; their frequency distribution is shown in Fig. 1.

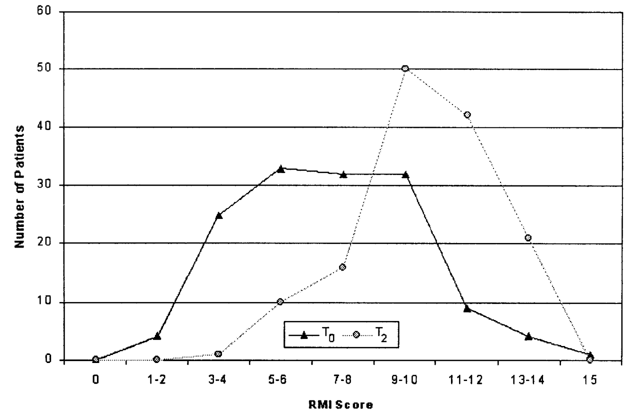

Fig. 1. Frequency distribution of the Rivermead Mobility Index scores, at the beginning (T_0) and at the end (T_2) of the prosthetic training.

The subgroup of 70 subjects showed the following values: at admission (T_0) RMI 8 (6–10), and motFIM 72 (66–79); at T_1 TWT 24.5 seconds (17–41); and at discharge (T_2) RMI 11 (10–13), motFIM 84 (81–87), and TWT 22.5 seconds (13–38).

The Cronbach's alpha of the RMI was 0.85, and the item-to-total correlation coefficients r_{pb} ranged, at T_0, 0.33–0.74 ($p < 0.0001$) for the items considered; the correlation regarding items 1, 3, 14 and 15 was not calculated, due to their extremely low (or non-existent) variability (mode $\geq$98%).

The correlation of RMI score with motFIM was $r_s = 0.83$ ($p < 0.0001$) at T_0, and $r_s = 0.69$ ($p < 0.0001$) at T_2, and that with TWT $r_s = -0.70$ ($p < 0.0001$) at T_2.

The median of the change scores was 3 (interquartile range 2–6); 61% of the patients ($n = 85$) showed a clinically significant improvement in the RMI (differences of 3 or more points between T_2 and T_0). The correlation of the change scores for the RMI with changes in motFIM, over the testing period, was $r_s = 0.75$ ($p < 0.0001$). The effect size was 1.35.

Fig. 2 shows the percentage of patients who passed each RMI item (scoring 1 point) at the 2 examination times, and allows an easy visual inspection of the hierarchical order and spacing of items along the underlying construct. The CR and CS ranged 0.91–0.94 and 0.38–0.62, respectively.

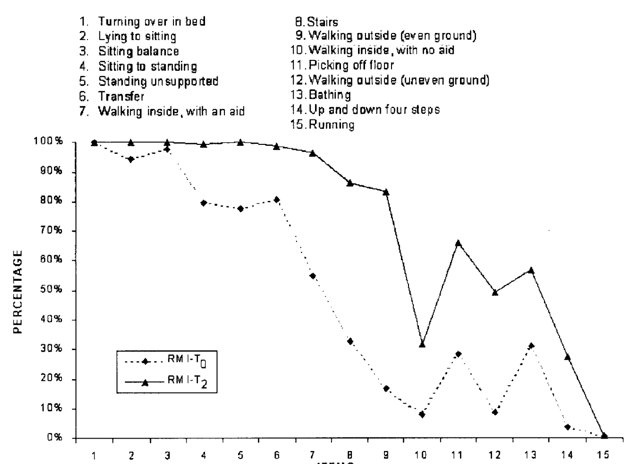

Fig. 2. Percentage of patients (n = 140) who passed each item of the Rivermead Mobility Index, at the beginning (T_0) and at the end (T_2) of the prosthetic training.

DISCUSSION

The present study shows that the RMI—when applied to LLA undergoing rehabilitation and prosthetic training—fulfils a series of classic psychometric requirements (for internal consistency, construct validity and responsiveness) needed for recommending an instrument in routine practice (11, 17).

The Cronbach alpha and the item-to-total statistics indicate an adequate homogeneity of the RMI items (i.e. the items tap different aspects of the same attribute). In fact, the criteria of alpha values above 0.70 and of the correlation of each item with the total score above 0.20 are met. However, all these values are considerably lower than those reported in stroke populations (18), indicating that the instrument, when applied in LLA, has a sufficient but not ideal internal consistency (Bland & Altman (19) stated that a minimum of 0.90 is desirable for the clinical application).

Regarding construct validity, we found (as expected) that the RMI showed a good correlation with both motFIM (independence in basic daily activities), and TWT (short-distance performance in walking), 2 measures assumed to be valid indicators of dimensions related to mobility. A similar correlation (ranging from 0.65 to 0.81) was previously reported between the RMI and the Barthel index, another well-known disability rating scale, in 144 above-knee amputees for vascular disease (9). Furthermore, some studies in patients with neurological diseases (2–8) showed a high correlation of the RMI with other mobility scales, disability scales, balance scores and walking performances. Thus, the present findings confirm that, in LLA too, the RMI scores adequately reflect patients' ability to move their own body within the environment. Conversely, the test does not investigate or capture either the "quality" of the mobility performance or the overall patient ability to manage daily tasks, that may also be a consequence of adaptive behaviours (such as use of the wheelchair or taking advantage of someone's help).

The results also provide evidence in support of the responsiveness of the RMI in the patients studied (14, pp. 103–105): most (61%) of the patients improved by 3 or more RMI points, the criterion of clinically significant improvement set by a previous study (7) and based on considerations about the reliability of the measure (2); and the relationship between the change in scores of the RMI and motFIM was high (0.75), which suggests that the RMI scores parallel the recovery of functional status.

Furthermore, the RMI is generally easy and quick to manage, demonstrating a low burden for the respondent and ease of administration and processing. Guidelines for its use in LLA would be needed. We adopted an Italian version of the RMI (18) produced—without any semantic difficulties being found—according to validated procedures of cross-cultural adaptation (20) and scored the items without considering the use of the lower limb prosthesis as an "aid" or "support" (see item 5 and 10).

Considering appropriateness and precision (2 additional criteria for selecting an outcome measure) (11), it is crucial to inspect the subject matter of the RMI in relation to its intended purpose, judging among others the face validity ("what an item appears to measure based on its manifest content"), the content validity ("whether the scale comprehensively samples all the relevant domains of interest, rather than unrelated and unintended aspects") and the number and accuracy of the distinctions made by the instrument.

The index covers with its 15 items the range from "completely bedridden" to "fully mobile with a prosthesis without aid". Accordingly, the RMI values of our sample ranged 1–15 at admission and 5–15 at discharge, with only 1 subject attaining the top score. The easiest tasks ("turning over in bed", "lying to sitting" and "sitting balance") capture the lowest level of mobility in the more frail LLA, whereas the questions about demanding tasks such as "go up and down four steps with no rail" and "running" investigate 2 very high achievements for indoors and outdoors prosthetic use, respectively. The choice of these items, originally selected for stroke patients, is in line with the range of measurement of the different mobility scales for LLA discussed in a recent review (1) and the most difficult ones cover a range of activities that seem able to adequately investigate the prosthetic mobility (21).

The issue of precision can be raised in relation to the scalability and the link between the difficulty of the single items and the "true" distribution of what is being measured. As regards scalability (16), the RMI does not meet the Guttman scaling criteria (coefficient of reproducibility >0.9, coefficient of scalability >0.65), even when the item columns have been rearranged to minimize the scaling errors. As a consequence, the RMI cannot be considered in LLA as a "hierarchical" scale in which any single item subsumes all those below (the ideal deterministic model is rarely if ever achieved in clinical practice, particularly in assessing functional loss) (17, p. 44), but only as a summated index with ordinal properties.

As regards the distribution of items according to their difficulty, Fig. 2 shows that the scale presents a similar level of difficulty in the first 3 items (as well as, but to a lesser extent, in items 4–6), an extremely skewed distribution of many item responses, and an uneven spacing of items along the continuum of the underlying construct which compromises the ability to calculate meaningful change scores (11, 22). Moreover, the rank order based on the percentage of passes for items shows in LLA an order of difficulty different from that of the original one (2) and that in stroke patients (18).

Improvements of the scaling properties of the RMI could be made through: (*i*) the deletion of some items (e.g. a couple of those representing similar and very easy levels of difficulty), and the addition of new items to fill in gaps along the unidimensional continuum; (*ii*) the expansion of the number of response categories to 3–5 ordinal levels (e.g. examining in more detail the need for external help and the use of aids). This latter modification would probably increase the precision of the scale, currently limited by the binary nature of the responses (yes/no) which does not provide information on the degree of difficulty in performing each task—an aspect that could be important to distinguish during rehabilitation training. A similar change was proposed and tested in 2 previous studies on neurological patients (8, 23).

Both these options require further psychometric testing of the scale so modified, but—as previously stated—the risk of misinference in interpretation of the total score and change score is already inherent in the original RMI, due to its ordinal scaling (24, 25). For this reason, Rasch measurement models (25, 26) seem warranted in order to further clarify some psychometric properties of the scale (e.g. unidimensionality, hierarchy, and interval location of items), by examining patterns of individuals' performances on the range of items and patterns of item difficulty. Rasch analysis would also be valuable in developing an equal-interval measure from raw rating scores. Furthermore, before introducing any modifications to the original scale, we think—for reasons of parsimony—that there is a need for more head-to-head studies comparing a number of mobility measures for LLA, including the RMI, the Locomotor Capabilities Index (21, 27), the Houghton Scale (21), and others (1). In addition, further studies are required in order to analyse the possible differences in difficulty profile of the RMI according to clinical and demographic factors (such as level and aetiology of amputation, age, gender, and so on).

CONCLUSION

In conclusion, the present findings indicate that, according to classic psychometric requirements, the RMI is a measure with acceptable levels of internal consistency, construct validity and responsiveness, when assessing the overall body mobility in LLA. However, further steps should be considered to improve its item selection, response format and scaling properties. Hence, the index seems at present more useful for epidemiological studies than for the everyday clinical application in single patients, where the identification of specific areas for treatment and a precise monitoring of the intervention results are required.

REFERENCES

1. Rommers GM, Vos LD, Groothoff JW, Eisma WH. Mobility of people with lower limb amputations: scales and questionnaires: a review. Clin Rehabil 2001; 15: 92–102.
2. Collen FM, Wade DT, Robb GF, Bradshaw CM. The Rivermead Mobility Index: a further development of the Rivermead Motor Assessment. Int Disabil Stud 1991; 13: 50–54.
3. Forlander DA, Bohannon RW. Rivermead Mobility Index: a brief review of research to date. Clin Rehabil 1999; 13: 97–100.
4. Vaney C, Blaurock H, Gatter B, Meiseds C. Assessing mobility in multiple sclerosis using the Rivermead Mobility Index and gait speed. Clin Rehabil 1996; 10: 216–226.
5. Molenaar DS, van Doorn PA, Vermeulen M. Pulsed high dose dexamethasone treatment in chronic inflammatory polyneuropathy: a pilot study J Neurol Neurosurg Psychiatry 1997; 62: 388–390.
6. Schindl MR, Forstner C, Kern H, Zipko HT, Rupp M, Zifko UA. Evaluation of a German version of the Rivermead Mobility Index (RMI) in acute and chronic stroke patients. Eur J Neurol 2000; 7: 523–528.
7. Hsieh CL, Hsueh IP, Mao HF. Validity and responsiveness of the Rivermead Mobility Index in stroke patients. Scand J Rehabil Med 2000; 32: 140–142.
8. Rossier P, Wade DT. Validity and reliability comparison of 4 mobility measures in patients presenting with neurologic impairment. Arch Phys Med Rehabil 2001; 82: 9–13.
9. Traballesi M, Brunelli S, Pratesi L, Pulcini M, Angioni C, Paolucci S. Prognostic factors in rehabilitation of above knee amputees for vascular diseases. Disabil Rehabil 1998; 20: 380–384.
10. Traballesi M, Amato S, Angioni C, et al. Efficacy and efficiency of the rehabilitation treatment of an amputee patient with vascular disease [in Italian]. Ann Ig 2002; 14: 27–35.
11. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. Health Technol Assess 1998; 2:1–74.
12. Guide for the Uniform Data Set for Medical Rehabilitation (including the FIM™ instrument), Version 5.1. Buffalo, NY 14214: State University of New York at Buffalo; 1997.
13. Collin C, Wade DT, Cochrane GM. Functional outcome of lower limb amputees with peripheral vascular disease. Clin Rehabil 1992; 6: 13–21.
14. Portney LG, Watkins MP. Foundations of clinical research. Applications to practice. 2nd edn. New Jersey: Prentice Hall Health; 2000.
15. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989; 27 (3 suppl): S178–189.
16. McIver JP, Carmines EG. Unidimensional scaling. London: Sage Publication Inc.; 1981: 47–51.
17. Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. 2nd edn. Oxford: Oxford University Press; 1995.
18. Franchignoni F, Tesio L, Benevolo E, Ottonello M. Psychometric properties of the Rivermead Mobility Index in Italian stroke rehabilitation inpatients. Clin Rehabil 2003 (in press).
19. Bland JM, Altman DG. Cronbach's alpha. BMJ 1997; 314: 572.
20. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993; 46: 1417–1432.
21. Miller WC, Deathe AB, Speechley M. Lower extremity prosthetic mobility: a comparison of 3 self-report scales. Arch Phys Med Rehabil 2001; 82: 1432–1440.
22. Stucki G, Daltroy L, Katz JN, Johannesson M, Liang MH. Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. J Clin Epidemiol 1996; 49: 711–717.
23. Lennon S, Johnson L. The modified Rivermead Mobility Index: validity and reliability. Disabil Rehabil 2000; 22: 833–839.
24. Johnston MV, Keith RA, Hinderer SR. Measurement standards for interdisciplinary medical rehabilitation. Arch Phys Med Rehabil 1992; 73 (12-S): S3–23.
25. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989; 70: 875–860.
26. Wright BD, Masters GN. Rating Scale Analysis. Chicago: Mesa Press; 1982.
27. Gauthier-Gagnon C, Grise MC, Lepage Y. The Locomotor Capabilities Index: content validity. J Rehabil Outcomes Meas 1998; 2: 40–46.