J Rehabil Med 2005; 37: 325–329

Taylor & Francis
Taylor & Francis Group

OUTCOME MEASUREMENT IN AUSTRALIAN REHABILITATION ENVIRONMENTS

Heather Douglas, Cheryl Swanson, Travis Gee and Nicholas Bellamy

From the Centre of National Research on Disability and Rehabilitation Medicine, Mayne Medical School, The University of Queensland, Herston, Queensland, Australia

Objective: **To determine the frequency and pattern of methods of outcome assessment used in Australian physical rehabilitation environments.**

Design: **Postal survey.**

Methods: **A questionnaire on service type, staffing, numbers of adults treated and outcome measures used for 7 conditions related to injury and road trauma as well as stroke and neuromuscular disorders was sent to 973 services providing adult physical rehabilitation treatment.**

Results: **Questionnaires were completed by 440 service providers for a response rate of 45%, similar to that reported in a recent European survey reported in this journal. A small number of measures were reported as in use by most respondents, while a large number of measures were used by a few respondents. Measures of physical changes were used more frequently than those of generic well-being or quality of life. Ease of use and reporting to other professionals were cited as the most important reasons in selection of outcome measures.**

Conclusion: **This Australian-wide survey detected considerable heterogeneity in outcome measurement procedures used in rehabilitation environments. While the goal of measurement may vary between providers and differ between conditions, the results highlight opportunities for harmonization, benchmarking and measurement of health-related quality of life.**

Key words: Australia, rehabilitation, outcome measures, change, standardization, benchmarking.

J Rehabil Med 2005; 37: 325–329

Correspondence address: Cheryl Swanson, Centre of National Research on Disability and Rehabilitation Medicine, Mayne Medical School, The University of Queensland, Herston, Queensland 4006, Australia. E-mail: c.swanson@uq.edu.au

Submitted June 7, 2004; accepted February 24, 2005

INTRODUCTION

Outcome measurement is primarily used in the evaluation of treatment efficacy and the quantification of rehabilitation results. The assessment of interventions and their outcomes allows providers to evaluate patterns of response and provides a mechanism for maximizing opportunities for facilitating improved management and outcomes (1). Outcome measurement is increasingly becoming an important component of decision-making and quality assurance in rehabilitation service delivery. An additional goal of rehabilitation service delivery, as well as health service delivery in general, is benchmarking to enable organizations to make comparisons against other similar services. Such a goal, however, necessitates a standardization of measurements that can be used across settings and conditions which, in turn, requires the use of measures that best identify the most appropriate and effective treatments and strategies. The measures that satisfy these criteria are yet to be determined, although some measures, such as the Functional Independence Measure (FIM^{TM})(2), the Barthel Index (both original and modified (3, 4)), and the SF-36 (5) are regarded as suitable, given their use in a wide variety of settings and conditions and, in the case of the FIM^{TM} (2) and SF-36 (5), their association with established databank facilities (Uniform Data System (www.udsmr.org) and QualityMetrics (www.sf36.org), respectively) in the USA.

Although assessment is a central component of rehabilitation practice, there has been little systematic examination of the precise outcome measurement procedures currently employed.

To address this question in Europe, Haigh et al. (6) conducted a postal survey reported in this journal in 2001. That report described a large number of measures being used, but a relatively small number being used frequently. We have conducted a similar survey in Australia to investigate and document patterns of outcome measurement in rehabilitation environments.

METHODS

A questionnaire modelled closely on that used by Haigh et al. (6) was developed for Australian rehabilitation environments. Six conditions, low back pain, spinal cord lesions, traumatic brain injuries, lower limb amputations, stroke, and neuromuscular disorders, as well as their definitions, were retained from the original study, while 3 conditions in the original questionnaire, hip and knee replacement, multiple sclerosis, rheumatoid arthritis, were replaced with whiplash, burns and upper limb amputation to reflect interests associated with the Motor Accident Insurance Commission, the funding body for our centre. Lists of potential outcome measures provided for each condition were derived from the relevant literature. Minor changes in phraseology and format were made to the original questionnaire and questions on reasons for selecting and using outcome measures were included. The questionnaire was then tested in 2 Australian States (New South Wales and Queensland) with 16 experienced practitioners and academics representing the major rehabilitation professions in public and private settings, who recommended minor changes to reflect locally used terminology and outcome measures.

Respondents were asked to provide information on staff numbers, funding sources and estimations of the numbers of adults treated

 ISSN 1650-1977
DOI 10.1080/16501970510035089

annually by the service. Respondents were asked to provide information on their organization and service type as acute (intensive), sub-acute (maintenance), slow stream (review/monitoring) or other. Respondents were asked to identify and rate the importance of both the reasons for their use of and the preferred features of outcome measures. A separate page for each of the 9 conditions requested details of the estimated number of adults treated annually, the specific outcome measures used, the estimated number of annual assessments made with each outcome measure and the name of any other measure used that was not included in the list provided. The categories used in this study were the same as those used by Haigh et al. (6). The final questionnaire page offered space for comments. The lower and upper bounds for the numbers of adults treated and assessments performed in this study were determined by multiplying the lower and upper bounds of each category (e.g. 1–25, 26–50, etc.) respectively by the number of respondents selecting that category and summing those figures to obtain the total lower bound and total upper bound in each instance. Descriptive statistics, frequencies and cross-tabulations were performed using SPSS (7).

The survey process

Potential participants were identified from State and Territory listings of hospitals, listings of rehabilitation services in the Yellow Pages®, and the non-retired membership of the Australasian Faculty of Rehabilitation Medicine. This process identified 1748 possible participants who represented all States and Territories and included private providers (55%), public hospitals (29%), rehabilitation physicians (13%) and private hospitals (3%).

Of these potential participants, 775 were excluded because of duplicate listings or provision of physical rehabilitation treatment service that did not include adult patients. This figure also included 151 Commonwealth Rehabilitation Services providing case management but no specified treatment and 95 surveys that were returned with addresses unknown. Of the resulting 973 possible participants, a total of 448 informative surveys were received. Of these respondents, 8 provided organizational information only and were not included in the analysis, resulting in 440 who returned completed surveys. The proportion of respondents ranged from 36% (Northern Territory) to 70% (Australian Capital Territory) and the overall response rate was 45%. Non-respondents were similar to respondents in terms of type of organization; i.e. private providers (267 (55%)), public facilities (149 (30%)) and medical rehabilitation specialists (109 (13%)).

RESULTS

The estimated number of adults treated in rehabilitation each year was reported as between 70,491 and 92,660, of whom 19,569–25,622 were inpatients; 40,288–50,826 were outpatients or community-based clients; 5839–10,164 were home-based clients; and 4795–6048 were described as "Others". Outpatient or community-based settings provided 28% of treatments, rehabilitation services provided 25% and public hospitals provided 22% and some respondents provided services in more than 1 setting. Subacute or maintenance rehabilitation was provided by 36% of respondents, with acute services provided by 27%, slow stream by 27% and other kinds of services provided by 10%. A variety of staff were employed by these service providers, including 2936 full-time equivalent (FTE) nursing staff, 984 FTE physiotherapists, 828 FTE occupational therapists and 326 FTE physicians.

The majority of respondents reported providing treatment for low back pain (74.5%), followed by whiplash (51.8%), strokes (47.7%) and neuromuscular disorders (45.7%), while relatively few respondents reported providing treatment for burns (15%) or upper limb amputations (13%).

Table I. *Most frequently cited outcome measures: low back pain, whiplash, spinal cord lesions by number of services (*n*) and estimated number of assessments*

Outcome measure	*n**	Assessments annually (min.–max.)
Low back pain	328	
Range of Movement	279	30859–56779
Visual Analogue scale	246	21327–43115
Muscle Function Testing	220	21253–41019
Oswestry Low Back Pain	115	4550–12950
FIM™	88	1371–5300
McGill Pain (short form)	81	2618–7251
Roland & Morris Disability	59	1094–4950
SF-36	49	841–3050
Quebec Back Pain Disability	42	980–3250
Canadian Occupational Performance Measure	36	874–2850
Whiplash	228	
Range of Movement	195	9920–22509
Visual Analogue Scale	156	6137–15954
Muscle Function Testing	143	7128–16156
Neck Disability Index	61	2202–5951
McGill Pain (short form)	52	796–3650
FIM™	32	424–1850
Fear Avoidance Belief Scale	27	570–2100
General Health Questionnaire	23	868–2600
SF-36	22	216–1250
Spinal cord lesion	119	
Manual Muscle Testing	93	2078–6252
FIM™	73	1408–4601
ASIA Motor Score	30	978–2800
Visual Analogue Scale	34	229–1700
Lung Vital Capacity	33	526–2150
ASIA Light touch	25	774–2150
Modified Ashworth Scale	25	221–1600
ASIA Pinprick	24	773–2100
ASIA Impairment Scale	23	571–1700
Barthel Index (modified)	23	120–1150

* *n* = number of respondents.

The most frequent reason cited as an influence for use of outcome measures was the requirement to report or refer to other health professionals, which 91% of respondents rated as "somewhat important" or "very important". Ease of use, reliability and validity of measures were the features for a suitable outcome measure most frequently cited as important by respondents, 99% of whom rated these features as "somewhat important" or "very important". "Detects change in quality of life" was regarded as slightly more important a feature than "Detects change in physical function", cited as "somewhat important" or "very important" by 97.3% and 96.2%, respectively. "Used by similar services" was cited by 97.3% of respondents as "somewhat important" or "very important".

The most frequently reported outcome measurements for each condition are presented in Tables I–III. Respondents also reported using a number of measures beyond those listed in the questionnaire for each condition, ranging from at least 10 other measures in burn injury rehabilitation to over 100 other measures in stroke and low back pain rehabilitation.

Table II. *Most frequently cited outcome measures: neuromuscular disorders, traumatic brain injury, stroke by number of services (*n*) and estimated number of assessments*

Outcome measure	*n**	Assessments annually (min.–max.)
Neuromuscular disorders	201	
Range of Movement	171	3843–11701
Muscle Function Testing	151	2871–9851
FIM™	96	1731–6700
Timed up and go	89	1125–5300
Dynamometry	72	1160–4201
Berg Balance Scale	56	899–3700
Barthel Index (modified)	45	938–3500
Quantitative Sensory Tests	39	837–2950
Lung Vital Capacity	38	80–1550
Modified Ashworth Spasticity	32	427–2100
Canadian Occupational Performance Measure	30	74–1350
9 Hole Peg Test	27	437–1550
Traumatic brain injury	146	
Mini Mental State Exam	79	814–4151
FIM™	74	1409–4701
Westmead PTA Score	65	846–3201
Glasgow Coma Scale	61	942–2251
Rivermead Behavioural Memory	39	186–2000
Barthel Index (modified)	34	129–1600
Modified Ashworth Scale	23	17–850
Berg Balance Scale	25	18–900
Barthel Index (original)	12	13–650
Brisbane Perceptual Screening	11	59–500
Stroke	210	
Mini Mental State Examination	143	4613–12501
FIM™	108	4289–10401
Berg Balance Scale	67	1606–5150
Barthel Index (modified)	59	1850–5450
Western Aphasia Battery	56	895–3350
Glasgow Coma Scale	66	895–3550
Rivermead Behavioural Memory	43	636–2550
Modified Ashworth Scale	37	730–2550
Rivermead Perceptual Test	34	477–1950
Beck Depression Inventory	23	269–1300
Brisbane Perceptual Screening	18	263–950
SF-36	17	215–1050

* *n* = number of respondents.

Table III. *Most frequently cited outcome measures: Upper limb amputation, lower limb amputation and burns by number of services (*n*) and estimated number of assessments*

Outcome measure	*n**	Assessments annually (min.–max.)
Upper limb amputation	57	
FIM™	16	313–1100
McGill Pain Short Form	11	61–610
Canadian Occupational Performance Measure	11	9–450
Purdue Pegboard Test	10	9–450
Preston Pinch Gauge	9	58–450
Minnesota Rate of Manipulation	6	6–300
Lower limb amputation	129	
Range of Movement	115	743–5550
Manual Muscle Testing	100	784–5150
FIM™	74	713–4100
10 Metre Walk	64	454–3250
Timed Up and Go	52	193–2300
Visual Analogue Scale	43	136–1900
Barthel Index (modified)	32	176–1500
Canadian Occupational Performance Measure	21	118–1000
Burns	66	
Range of Movement	56	52–2600
Physical examination	55	51–2550
Dynamometry	23	23–1150
2-point discrimination	20	19–950
FIM™	19	566–1301
Barthel – modified	12	11–1301
Canadian Occupational Performance Measure	8	6–300

* *n* = number of respondents.

The 10 measures most likely to be reported were compared across the conditions with regard to numbers of respondents reporting their use (Table IV). Range of movement (ROM) (8) and manual muscle testing (MMT) (9), reported in use in 5 of the 9 conditions, were used by at least 85% and 62%, respectively, of those respondents. Although the FIM (2) was reported in use for all but 1 of the conditions (8/9), its level of usage in each condition ranged from 27% to 61% by respondents. The SF-36 (5) was not widely used amongst conditions (3/9), nor was its level of usage high, ranging from only 8% for stroke to 15% for low back pain.

DISCUSSION

This national study is an important development in describing the use of outcome measurement procedures in rehabilitation environments in Australia. A total of 973 Australian public and private service providers were located and a response rate of 45% achieved, which was comparable to the 48% reported for the European survey (6). Responders did not differ from non-responders in their organizational affiliations, suggesting that the results are likely to be generalizable to the Australian rehabilitation community as a whole.

The results of this Australian survey show many similarities to those in the European survey (6). The 3 most frequently reported measures within each condition were the same in both regions for stroke, traumatic brain injury and lower limb amputation although differing sometimes in order. The 2 most frequently reported measures in low back pain (ROM (8) and Visual Analog Scale (10)) and spinal cord injury (MMT (9) and FIM™ (2)) were the same in both surveys, while the third most frequently reported measure differed between Australia and Europe for both low back pain (MMT (9) and McGill Pain (short form) (11), respectively) and spinal cord lesion (ASIA motor score (12) and Modified Ashworth Scale (13), respectively).

Both surveys found considerable variability in outcome assessment procedures across rehabilitation services, and substantial variation in the measures used for the conditions studied. In general, rehabilitation services in both Europe and Australia assessed change in physical status, and the major level of assessment was at the impairment level using ROM (8) and

Table IV. *Frequency (percentages) of use of measures by services (n (%)) across conditions*

	Condition								
Measure	LBP *n* (%)	WAD *n* (%)	SCL *n* (%)	NMD *n* (%)	TBI *n* (%)	Stroke *n* (%)	ULA *n* (%)	LLA *n* (%)	Burns *n* (%)
ROM	279 (85.1)	195 (85.5)		171 (85.1)				115 (89.1)	56 (84.8)
MMT	220 (67.1)	143 (62.7)	93 (78.2)	151 (75.1)				100 (77.5)	
FIM^{TM}	88 (26.8)		73 (61.3)	96 (47.8)	74 (50.7)	108 (51.4)	16 (28.1)	74 (57.3)	19 (28.8)
MBI			23 (19.3)	45 (22.4)	34 (23.3)	59 (28.1)		32 (24.8)	12 (18.2)
COPM	36 (11.0)			30 (14.9)			11 (77.5)	21 (16.3)	8 (12.1)
SF-36	49 (14.9)	22 (9.6)				17 (8.1)			
BBS				56 (27.9)	25 (17.1)	67 (31.9)			
MP-S	81 (24.7)	52 (22.8)					11 (77.5)		
VAS	216 (65.9)	156 (68.4)	34 (28.6)					43 (33.3)	
RBM					39 (26.7)	43 (20.5)			

* ROM = Range of Movement; MMT = Manual Muscle Testing/Muscle Function testing; FIM = Functional Independence Measure; MBI = Modified Barthel Index (+ original Barthel); COPM = Canadian Occupational Performance Measure; SF-36 = MOS Short Form 36 Questionnaire; BBS = Berg Balance Scale; MP-S = McGill Pain (short form); VAS = Visual Analog Scale; RBM = Rivermead Behavioural Memory; LBP = low back pain; WAD = whiplash-associated disorder; SCL = spinal cord lesion; NMD = neuromuscular disease; TBI = traumatic brain injury; ULA = upper limb amputation; LLA = lower limb amputation.

MMT (9) measures. A wide range of pain scales was used, but pain was not measured frequently. Disability, impact of events and integration scales were rarely reported and little assessment made of psychosocial adjustment or mental health status using depression or anxiety scales by either European or Australian services. The FIM^{TM} (2) was the measure used most commonly to assess function in both regions, with its use reported in 8 of 9 conditions in Australia and 7 of 9 conditions in Europe (6). Both surveys found a low use of the SF-36 (5) in practice in direct contrast to its frequent use in research studies (14).

Many assessment instruments exist and no single measure is accepted as providing all essential information for the conditions studied across all service types in both Australia and Europe. A pattern has emerged in both regions of many outcome measures being used with a relatively small number of measures used frequently, and of numerous measures cited in addition to those provided in the survey questionnaire. It appears, therefore, that the diversity relates to factors other than a geographic location.

Reasons for using and the preferred features of outcome measures were sections newly included in this Australian survey (c.f. the European survey), and the relative frequencies of the responses have provided important insights into the process of selecting measures even though the respondents were not asked to rank the individual items. Reporting or referring to other professionals, reporting to insurance companies, reporting for litigation and reporting to healthcare administration were cited as "very important" or "somewhat important" by at least 75% of respondents and suggests that external communication needs to play an important role in the use of outcome measures. "Easy to use" was the feature most often cited as important, suggesting that efficiency is likely to be a major consideration for service providers. Interest in standardization or benchmarking appeared to be substantial, as suggested by the item "Used by similar services", with 88.5% rating it as "very important" or "somewhat important". The item "Detects change in quality of life" was rated "very important" or "somewhat important" by 97.3% of respondents, although few specific quality of life measures were reported in use to any extent amongst any of the conditions. Reliably validated measures of health-related quality of life are available and should be more widely used.

Rehabilitation environments are changing rapidly and are susceptible to multiple influences, some of which, e.g. potential conflicts between the needs of different insurers vs needs in benchmarking, may be changing as well. Identifying and understanding these influences and changes can enhance assessment practices for the benefit of the recipients of that rehabilitation. The standardization of assessment practices will contribute greatly to the harmonization of outcome measurement procedures in rapidly changing rehabilitation environments. To this end, the establishment of the Australian Rehabilitation Outcomes Centre (AROC) has been an important step towards the goals of providing a national "data bureau" for data on the performance of rehabilitation services in this country as well as a national "benchmarking centre" for medical rehabilitation.

Benchmarking at an international level may be a desirable goal, however, many issues would need to be addressed to ensure its success. Such issues would include methodological and jurisdictional differences in practice, access and patient management, as well as demographic and cultural differences and available and appropriate translations of assessments. Recent research in cross-cultural issues in impairment and assessment using such widely used assessments as the FIM^{TM} (2) has drawn attention to the challenges inherent in achieving cross-cultural validity (15).

A better understanding of the accepted principles and information needs in the industry, both in Australia and Europe, would inform future planning and development of outcome measurement procedures. Although a core set of outcome measures has yet to be identified, this survey forms the basis for future study of the opportunities for, and benefits of, standardization of measuring outcomes. Such standardization would

contribute to improved patient and client outcomes and facilitate comparisons across different service types and conditions both nationally and internationally.

ACKNOWLEDGEMENTS

We thank Professor Alan Tennant from University of Leeds for providing the original survey questionnaire. We also thank the individuals who participated in the pilot survey and staff of CONROD for help in the design of the questionnaire and conduct of this study. This study was funded as part of the core activities of CONROD.

REFERENCES

1. Cohen ME, Marino RJ. The tools of disability outcomes research: functional status measures. Arch Phys Med Rehabil 2000; 81 (2 Suppl): S21–S29.
2. Keith RA, Granger CV, Hamilton BB, Sherwin FS. The functional independence measure: a new tool for rehabilitation. In: Eisenberg MG, Grzesniak RC, eds. Advances in clinical rehabilitation. New York: Spring Publishing; 1987, vol. 1. pp. 6–18.
3. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Maryland State Med J 1965; 14: 61–65.
4. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989; 42: 703–709.
5. Ware JE, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30: 473–483.
6. Haigh R, Tennant A, Biering-Sorensen F, et al. The use of outcome measures in physical medicine and rehabilitation within Europe. J Rehabil Med 2001; 33: 273–278.
7. SPSS for Windows, Rel. 11.0.1. 2001. Chicago: SPSS Inc.
8. Gajdosik RL, Bohannon RW. Clinical measurement of range of motion. Review of goniometry emphasizing reliability and validity. Phys Ther 1987; 67: 1867–1872.
9. Gonnella C. The manual muscle test in the patient's evaluation and program for treatment. Phys Ther Rev 1954; 34: 16–18.
10. Joyce CR, Zutshi DW, Hrubes V, Mason RM. Comparison of fixed interval and visual analogue scales for rating chronic pain. Eur J Clin Pharmacol 1975; 8: 415–420.
11. Melzack R. The short-form McGill Pain Questionnaire. Pain 1987; 30: 191–197.
12. American Spinal Injury Association. International Standards for Neurological Classification for Spinal Injury. Chicago: American Spinal Injury Association; 2000.
13. Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. Phys Ther 1987; 67: 206–207.
14. Ware JE, Jr. SF-36 health survey update. Spine 2000; 25: 3130–3139.
15. Tennant A, Penta M, Tesio L, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: The PRO-ESOR Project. Med Care 2004; 42 (1 Suppl): 37–48.