Submission to *Neurorehabilitation* - 2015

Palmer, C., Kneebone, I. I., Strauss, C. & Jones, M. (2016). Using reliability of change analysis to evaluate post-acute neuro-rehabilitation. NeuroRehabilitation. 38, 93–98. doi:10.3233/NRE-151300

Using reliability of change analysis to evaluate post-acute neuro-rehabilitation

Christina Palmer* [1], Ian I. Kneebone [2, 3], Clara Strauss [4, 5] and Anna-Marie Jones [4]

[1] University of Surrey, School of Psychology, Guildford, UK.

[2] Virgin Care, Department of Psychology, Haslemere, UK.

[3] University of Technology Sydney, Discipline of Clinical Psychology, Graduate School of Health, Sydney , Australia.

[4] Sussex Partnership NHS Foundation Trust, Research and Development Department, Worthing, UK.

[5] University of Sussex, School of Psychology, Brighton, UK.

* Note Dr Palmer is now with Sussex Partnership NHS Foundation Trust, Worthing, UK.

Address for correspondence: Professor Ian Kneebone
University of Technology Sydney,
Graduate School of Health
PO Box 123
Broadway NSW 2007
Australia
Ph: +61 2 9514 1448
Email: ian.kneebone@uts.edu.au

Abstract

Background: It is important to evaluate change in order to re-assure commissioners, staff and patients of the effectiveness of interventions, but also in order to identify areas for improvement. Objective: To consider whether analysis of improvement at the level of the individual, taking into account measurement error, may offer a further valuable way to assess change and inform service development over considering change at the group level in a post-acute neuro-rehabilitation unit. Method: Pre and post intervention Scores on the FIM+FAM Full Scale and Cognitive and Motor subscales were considered for eighteen patients aged between 35 and 81 with mixed diagnoses who attended a post-acute inpatient neuro-rehabilitation unit for treatment. Results: Statistically significant improvements were achieved on the FIM+FAM Full Scale and Cognitive and Motor subscales in a whole group analysis. Reliable change analyses for each patient within each subscale however identified only half of the sample achieved reliable improvement within the Motor domain and just one person within the Cognitive domain (5.6%). Conclusions: Findings are consistent with the emphasis of the rehabilitation unit on physical/motor function, and unsurprising as many of those assessed had multiple sclerosis, an often deteriorative condition. Use of reliable change analysis allowed a more detailed understanding of intervention impact, potentially identifying what services reliably work for whom, thereby informing future planning.

Keywords: neurorehabilitation, reliable change analysis, neurodisability, neurological illness

Running head: Reliability of change analysis in neurorehabilitation

Introduction

Inpatient neurological rehabilitation offers a coordinated multidisciplinary (MD) approach to treating a wide range of neurological conditions. The aim of treatment is to maximise potential for physical, cognitive, social and psychological functioning to enhance people's ability to participate in society and their quality of life (Turner-Stokes, 2010). This fits with the aims of the National Service Framework (Department of Health, 2005) which emphasises the need to support people with long-term neurological conditions to achieve the greatest possible level of independence and social inclusion.

There has been a recent shift in health services away from process targets and onto measuring clinically relevant outcomes (Department of Health, 2010). Measuring clinical outcomes within rehabilitation programmes provides a means of determining the "extent to which the interventions have achieved their aims" (Royal College of Physicians & British Society of Rehabilitation Medicine, 2003 p. 22). This provides valuable information for commissioners, staff and patients, which can also be used to guide service development and drive up the standards of care being delivered (Department of Health, 2008).

Despite some of the inherent challenges associated with evaluating a vastly heterogeneous patient group presenting with a diffuse and diverse range of deficits, there is an emerging evidence base which suggests that neurological rehabilitation is an effective intervention for patients diagnosed with a range of neurological impairments. These include acquired brain injury (ABI), Traumatic Brain Injury (TBI), stroke, multiple sclerosis (MS), spinal cord injuries, neuromuscular conditions and other brain pathology (Prvu Bettger, & Stineman, 2007; Department of Health, 2005; Freeman, Hobart, Playford, Undy & Thompson, 2005; Turner-Stokes, 2008). Studies investigating the differential effects of early versus delayed admission in TBI, stroke, and ABI populations have found that early post-acute rehabilitation is associated with better outcomes and reduced length of stay (Ashley, O'Shanick, & Kreber, 2009; Salter, Jutai, Hartley, Foley, Bhogal et al. 2006; Turner-Stokes, 2008).

In rehabilitations centres in the UK, functional outcome measures tend to be the most widely used measure of change (Skinner & Turner-Stokes, 2006). Typically within this field, change is measured at the group level. Whilst this clearly has its merits, relying solely on aggregate data analysis can be at the cost of valuable information regarding individual responses to treatment which might more fully inform practitioners and guide treatment planning and clinical practice. Reliable change methods were first popularised by Jacobson and Truax (1991) as means of determining the reliability of change within psychotherapy research, that is, change that is greater than would be expect than simply due to measurement error. Reliable change analysis was a novel approach as it provided a means of determining whether the magnitude of change for a given client was reliable. To date, while reliable change analysis has been utilised in neuro recovery/rehabilitation contexts such as to document outcomes post unruptured intracranial aneurysm (Towgood, Ogden & Mee, 2005), or cognitive changes following anterior temporal lobectomy (Dulay, Levin, York, Li, Mizrahi et al. (2009) and bilateral subthalamic nucleus deep brain stimulation (York, Dulay, Macias, Levin, Grossman et al. 2008) and in specific intervention research (e.g., Freedburg, Lynch III, & Ryan, 2011; Jenkins et al., 2006; Kesler, Sheau, Koovakkattu & Reiss, 2011), its potential usefulness in routine multi-disciplinary service evaluation has not been documented. It was expected that an evaluation of a neurorehabilitation service using reliable change analysis would provide additional information over and above that provided via traditional group analysis.

Method

Service setting

The ward evaluated was based in a community hospital offering services primarily to working age adults (18 - 65) who had experienced an acute neurological event such as stroke and people who had experienced changes in existing neurological conditions. Older adults (> 65) were accepted if they were considered able to participate and benefit from the intensive rehabilitation offered.

Sample

In total, 22 patients were admitted to the unit over a 6 month period. All patients for whom complete data were available were included in the study (*n* = 18, 81.8%). Patient's ages ranged between 35 and 81, with a mean age of 57.7 (*SD* 11.77) years. Age data was missing for one patient. The sample was comprised of 11 males (61%) and 7 females (39%). The length of admission ranged from 17 to 131 days with a mean of 45.28 days. The most common reason for admission was multiple sclerosis (*n* = 7, 38.9%) followed by stroke (*n* = 2, 11.1%). Other reasons for admission included myalgic encephalopathy, brain tumour, Cavernous hemangioma, Parkinson's disease, spinal cord infarct, spinal muscular atrophy, cervical myelopathy, cerebellar ataxia and subacromial decompression (all *n* = 1, 5.6%).

Measures

The FIM+FAM is comprised of the Functional Independence Measure (FIM): an 18 item scale of global disability and the Functional Assessment Measure (FAM) which adds twelve items which specifically address cognitive and psychosocial function (Hamilton, Granger, Sherwin, Sielenzny, & Tashman, 1987; Turner-Stokes, Nyein, Turner-Stokes & Gatehouse, 1999). The FIM+FAM is widely used in routine clinical practice in rehabilitation centres across the UK (Skinner & Turner-Stokes, 2006). Function is assessed by clinicians based on direct observations and scoring is undertaken by a multidisciplinary team. Ratings are made on a 7 point scale ranging from 1 (total assistance) to 7 (complete independence). Ratings are made at admission within 10 working days and during the final week before discharge. Factor analysis has identified that items fall within two subscales which reflect cognitive and motor domains (Turner-Stokes & Seigert, 2013). The FIM+FAM has been found to be a valid and reliable measurement of neurologic disability though with some variability in the responsiveness to change across items (Hobart et al., 2001) ; the motor scale appears to demonstrate greater responsiveness to change than the cognitive scale. Reliability ratings for cognitive items have also been found to be lower than motor items which may be attributable to the abstract nature of items (Alcott, Dixon & Swann, 1997). Findings from a study (Hawley, Taylor, Hellawell & Pentland, 1999) into the psychometric properties of the FIM + FAM justify treating the "raw scores as good and useful approximations to points on interval scales of measurement and treating them arithmetically to characterise levels of functioning on the subscales" (Hawley et al., p. 754) though it is acknowledged homogeneity of a sample may be an issue in this.

Planned Analysis

This was a retrospective data analysis evaluation of functional change scores between admission and discharge for patients accessing the unit over a twelve month period. Wilcoxon signed ranks tests were used to compare pre- and post-intervention scores for the motor and cognitive subscales as identified by Turner-Stokes and Siegert's (2013) factor analysis of the FIM+FAM.

Reliable change was calculated for each participant using the Leeds Reliable Change Index Calculator (Agostinis, Morley and Dowzer, 2008) which adheres to the method described by Jacobson and Truax (1991). This method involves subtracting the post treatment score from the pre-treatment score and dividing this by the standard error of the difference. The standard error of the difference describes the spread of distribution of change scores that would be expected if no actual change had occurred. In accordance with this method, reliable change index (RCI) scores of 1.96 or above were considered reliable ($p < .05$). Cronbach's apha (α) scores provided by Turner-Stokes and Siegert (2013) for the cognitive and motor subscale were used as parameters of the reliability of the measure.

Group change

Table 1 shows the medians and inter-quartile ranges for the full scale and subscale scores on the FIM+FAM measure at admission and at discharge. There were significant improvements at the group level on the full scale score ($Z(17) = -.3291$, $p < .001$, $r = 0.65$) and on the cognitive ($Z(17) = -2.873$, $p = .002$, $r = .048$) and motor subscales ($Z(17) = -3.297$, $p < .001$, $r = 0.55$).

Insert Table 1 about here

Reliable Change

Reliable change calculations were conducted for each patient for the motor and cognitive subscales. As can be seen in Table 2 with respect to the motor subscale, nine of the eighteen patients (50%) showed reliable improvement, as indicated by a change of 13.44 points or more and nine patients showed no reliable change (50%). With respect to the cognitive subscale, one patient (5.6%) demonstrated reliable improvement, one patient (5.6%) reliable deterioration (RCI reference score > 13.53), with 16 patients (88.8%) failing to demonstrate reliable change (Table 3).

Insert Table 2 and 3 about here

Discussion

This evaluation investigated whether changes occurred within functional domains following multidisciplinary treatment in a neurological rehabilitation unit. The findings demonstrated significant improvements at the group level in full scale scores at discharge as well as the cognitive and motor subscales. This improvement is generally consistent with the literature suggesting the efficacy of multi-disciplinary rehabilitation for patients with a range of neurological disorders and with the findings from a previous evaluation of the unit (Guerrier & Kneebone, 2004). Given that a large proportion (38.9%) of the sample had a diagnosis of MS, which is typically a deteriorative condition, these findings are impressive. A recent meta-analysis of cognitive rehabilitation in MS found that whilst there was some evidence of improvement following intervention when studies were evaluated individually, cumulatively the evidence for cognitive improvement was low (Rosti-Otajarvi & Hamalianen, 2011). This may be reflected in the discrepancy found between the degree of improvement on the motor and cognitive subscales. However, as we have demonstrated, significant group-level improvements are not necessarily indicative of reliable improvement at the individual level. Whilst the group results indicate improvement at the group level in terms of increased independence, analyses of individual patient data suggests that only half of the sample (50%) included in the evaluation demonstrated reliable improvements in scores within the motor domain and only one patient within the cognitive domain (5.6%). Within the cognitive domain, one patient (5.6%) demonstrated reliable deterioration. This is likely to reflect the emphasis of the unit at that time which was more focussed on physical goals and which would have influenced the nature of referrals received. Indeed, the unit has a sister service that prioritises cognitive rehabilitation and is the local head injury in-patient rehabilitation service. To some degree, this may also reflect factors associated with the FIM+FAM such as greater responsiveness to change across motor scale items or reliability issues due to the abstract nature of cognitive items (Alcott, Dixon & Swann, 1997).

As demonstrated by this study, the inclusion of reliable change methods provides another dimension to our analysis of the data gathered at the unit which has clear clinical relevance. Individual data analyses using this method demonstrated that whilst there was overall significant improvement following neurological rehabilitation in the setting evaluated, improvements were only reliable for half the patients in the motor areas of functioning and for only one patient in the cognitive domain. Clinically, this information is useful as it enables practitioners to gain insights regarding individual patient's responsiveness to treatment and provides a benchmark from which to assess the magnitude of change being observed. This information can be assimilated with other clinically relevant information in order to form a more comprehensive understanding of patients' progression through treatment. This enables practitioners to take a more person centred approach to treatment planning as well as providing information which can inform evaluation of patient centred goals. At a service level, this information not only provides commissioners with information regarding the effectiveness of interventions but also can help provide a more comprehensive picture regarding service provision. For example findings within this study reflected the emphasis within the unit on physical / motor functioning. In response to this, the service may want to consider ways of improving rates of cognitive functioning following treatment.

Alternatively, this may highlight issues associated with the relevance or adequacy of measures currently being used to capture clinical outcomes.

Limitations

There are various factors associated with the nature of neurorehabilitation research that make interpretations difficult. These include lack of control group, the vastly heterogeneous nature of the patient group, variations in the clinical manifestations of symptoms associated with neurological disorders or injury and difficulties accounting for factors such as spontaneous recovery.

In this particular study the lack of data regarding disease course for individual patients or duration between the acute and post acute phases of brain injury limits interpretations. This information is relevant given that the evidence suggests that early admission is associated with better outcomes (Ashley et al., 2009; Salter et al. 2006; Turner-Stokes, 2008). The relevance of data on disease course is particularly relevant to diagnoses such as MS where there are progressive and relapse remitting phases which will impact on functional ability.

Another limitation of the study is that outcome is measured solely in terms of clinician rated changes in functional outcome. This does not necessarily reflect the idiographic patient centred goals encompassing patient's views, cultural background and pre-morbid lifestyle that should be a key feature of rehabilitation (Royal College of Physician & British Society of Rehabilitation Medicine, 2003). Accordingly, standardised measures of functional outcome may not always be the most appropriate means of assessing meaningful change within this population, particularly when used alone. Similarly, reliable change analysis may not always be the most appropriate method for assessing the meaningful nature of changes in that it does not allow these factors to be accounted for.

A drawback of the FIM + FAM is that although it is designed for use in those who have suffered a brain-injury, this population is heterogeneous which notoriously adds complications when trying to interpret variations on the group level. The range of disability and need for assistance varies widely across the patient population and so the relative difficulties on a single item, such as bathing, will differ from patient to patient but this is not reflected in the scale.

The RCI aims to demonstrate clinical change beyond measurement error in such a way that an observed difference of a specific magnitude implies the same degree of clinical change regardless of the starting point of that patient. This clearly requires the summed raw scores to be equal-interval. However, it is likely that those with lower scores will have more complex medical conditions than those with higher scores and so for these patients the difference between a 1 (total assistance) and a 2 (maximal assistance) could potentially reflect a higher change in need than the difference between a 6 (modified independence) and a 7 (complete independence). Although we cannot demonstrate to what extent our current data fits the equal-interval description, there has been justification in treating the FIM+FAM scores in this way (Hawley et al., 1999) though admittedly on the basis of findings with a population potentially more homogeneous than that considered in this study. Meanwhile other investigators have argued that total scores for scales made up of many likert items can behave as interval scales (Carifio & Perla, 2008). The debate is ongoing (Perdices, 2004), more research is needed to externally validate reliable change in the raw FIM+FAM scores.

Declaration of interests

None.

References

Agostinis, A., Morley, S.J. & Dowzer, C.N. (2008). *The Leeds Reliable Change Indicator*. http://www.leeds.ac.uk/lihs/psychiatry/dclin/rcic.html [accessed 05.03.13].

Alcott, D., Dixon, K., & Swann, R. (1997). The reliability of the items of the Functional Assessment Measure (FAM): Differences in abstractness between FAM items. *Disability Rehabilitation*. 19, 9, 355-358.

Ashley, M., O'Shanick, G., & Kreber, L. (2009). Early vs. Late Treatment of Traumatic Brain Injury. Vienna, VA: Brain Injury Association of America.

Carifio, L., & Perla, R. (2008). Resolving the 50 year debate around using and misusing Likert scales. *Medical Education, 42*, 1150-1152.

Department of Health (2005). *National Service Framework for Long-Term Neurological Conditions*. London: Department of Health.

Department of Health (2008). *High Quality Care for All: NHS Next Stage Review Final Report*. London, HMSO.

Department of Health (2010). *Equity and excellence: Liberating the NHS* (White Paper). London: HMSO.

Dulay, M.F., Levin, H.S., York, M.K., Li, X., Mizrahi, E.M., Goldsmith, I., Verma, A., Goldman, A., Grossman, R.G., & Yoshor, D. (2009) Changes in individual and group spatial and verbal learning characteristics after anterior temporal lobectomy. *Epilepsia*. 50, 6, 1385-95.

Freedburg, M., Lynch III, J., & Ryan, J. P., (2011). Group environments and visuospatial attention in patients with TBI, *Scientia Discipulorum, 5*, 12-24.

Freeman, J.A., Hobart, J.C., Playford, E. D., Undy, B. & Thompson, A. J. (2005). Evaluating neurorehabilitation: Lessons from routine data collection. *Journal of Neurosurgery and Psychiatry*, 76, 723-728.

Guerrier, S., & Kneebone I. I. (2004). Pilot evaluation of a community hospital in-patient neurological rehabilitation programme. *International Journal of Therapy and Rehabilitation, 11*, 51-55.

Hamilton, B.B., Granger, C.V., Sherwin, F.S., Sielenzny, M., Tashman, M.J. (1987). A uniform national data system for medical rehabilitation. In M. J. Fuhrer (Ed). *Rehabilitation outcomes: Analysis and measurements*, pp. 137-147. Baltimore, MD: Brookes.

Hawley, C., Taylor, R., Hellawell, D., Pentland, B. (1999). Use of the functional assessment measure (FIM+FAM) in head injury rehabilitation: A psychometric analysis. *Journal of Neurology, Neurosurgery & Psychiatry, 67*, 749-754.

Hobart, J.C., Lamping, D.L., Freeman, J.A., Langdon, D.W., McLellan, D.L., Greenwood, R.J., & Thompson, A.J. (2001). Evidence-based measurement: Which disability scale for neurologic rehabilitation? *Neurology*, 57, 4, 639-44.

Jacobson, N.S. & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.

Jenkins, V., Shilling, V., Deutsch, G., l. Bloomfield, D., Morris, R., Allan, S… Winstanley, J. (2006). A 3-year prospective study of the effects of adjuvant treatments on cognition in women with early stage breast cancer. *British Journal of Cancer, 94*, 828–834.

Kesler, S. R., Sheau, K., Della Koovakkattu, D., & Reiss, A. L. (2011). Changes in frontal parietal activation and math skills performance following adaptive number sense training: Preliminary results from a pilot study. *Neuropsychological Rehabilitation, 21,* 433–454.

Perdices, M. (2014). Some thoughts about the suitability of the Reliable Change Index (RCI) for analysis of ordinal scale data. *Brain Impairment, 15*, 223-232.

Prvu Bettger, J.A. & Stineman, M.G. (2007). Effectiveness of multidisciplinary rehabilitation services in postacute care: State of the science. *Archives of Physical and Medical Rehabilitation*. 88, 1526-1534.

Rosti-Otajarvi, E. M., & Hamalainen, P. I. (2011). Neuropsychological rehabilitation for multiple sclerosis. *Cochrane Database of Systematic Reviews, 11*, CD009131.

Royal College of Physicians & British Society of Rehabilitation Medicine (2003). *Rehabilitation following acquired brain injury: National clinical guidelines*. London: Royal College of Physicians.

Salter, K., Jutai, J., Hartley, M., Foley, N., Bhogal, S., Bayona, N., & Teasell, R. (2006). Impact of early vs delayed admission to rehabilitation on functional outcomes in persons with stroke. Journal of Rehabilitation Medicine, 38,113-117.

Skinner, A. & Turner-Stokes, L. (2006). The use of standardised outcome measures in rehabilitation centres in the UK. *Clinical Rehabilitation, 20*, 609-615.

Towgood, K., Ogden, J. A., & Mee, E. (2005). Neurological, neuropsychological, and functional outcome following treatment for unruptured intracranial aneurysms. *Journal of the International Neuropsychological Society, 11*, 522–534.

Turner-Stokes, L. (2008). Evidence for the effectiveness of multidisciplinary rehabilitation following acquired brain injury: A synthesis of two systemic approaches. *Journal of Rehabilitation Medicine*, 40, 691-701.

Turner-Stokes, L. (2010). Levels of specialisation in rehabilitation services. http://www.bsrm.co.uk/ClinicalGuidance/Levels_of_specialisation_in_rehabilitation_services5.pdf [accessed 06.04.13].

Turner-Stokes, L., Nyein, K., Turner-stokes, T., & Gatehouse, C. (1999). The UK FIM+FAM: Development and evaluation. *Clinical Rehabilitation, 13*, 277-287.

Turner-Stokes, L. & Seigert, R.J. (2013). A comprehensive psychometric evaluation of the UK FIM+FAM. *Disability and Rehabilitation*, 35, 1885-1895.

York, M. K., Dulay, M., Macias, A., Levin, H. S., Grossman, R., Simpson, R., Jankovic, J. (2008). Cognitive declines following bilateral subthalamic nucleus deep brain stimulation for the treatment of Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*. 79, 789-795.

Table 1

Median and (Inter-Quartile Range) FIM+FAM scores on admission and discharge.

	Admission	Discharge
Full scale	171.5 (60.25)	195.0 (55.25)***
Cognitive	83.5 (17.75)	87.0 (16.50)**
Motor	94.0 (53.25)	108.0 (37.50)***

Note. *p < .05., **p < .01., ***p < .001.

Table 2

Calculations for reliable change on the FIM+FAM Motor subscale.

Participant	Admission length (days)	Reason for referral	Pre-treatment score	Post-treatment score	Difference	Reliable change met
1	91	Subacromial decompression	42	77	-35	yes
2	42	Chronic fatigue – Myalgic Encephalomyelitis	104	104	0	no
3	25	Multiple Sclerosis	108	108	0	no
4	29	Multiple Sclerosis	94	110	-16	yes
5	21	Stroke Haemorrhage	65	66	-1	no
6	35	Cerebella ataxia	95	122	-27	yes
7	133	Oligodendroma (grade 2)	32	34	-2	no
8	26	Spinal cord infarct	74	108	-34	yes
9	63	Multiple Sclerosis	40	66	-26	yes
10	76	Multiple Sclerosis	55	89	-34	yes
11	23	Multiple Sclerosis	118	121	-3	no
12	23	Spinal muscular atrophy	72	110	-38	yes
13	87	Stroke Haemorrhage	31	31	0	no
14	17	Cervical myelopathy	111	111	0	no
15	18	Parkinson	120	124	-4	no
16	27	Cavernous haemangioma	94	111	-17	yes
17	38	Multiple Sclerosis	98	114	-16	yes
18	41	Multiple Sclerosis	95	105	-10	no

Table 3

Calculations for reliable change on the FIM+FAM Cognitive subscale.

Participant	Admission length (days)	Reason for referral	Pre-treatment score	Post-treatment score	Difference	Reliable change met
1	91	Subacromial decompression	76	80	-4	no
2	42	Chronic fatigue – Myalgic Encephalomyelitis	81	84	-3	no
3	25	Multiple Sclerosis	87	88	-1	no
4	29	Multiple Sclerosis	83	86	-3	no
5	21	Stroke Haemorrhage	69	72	-3	no
6	35	Cerebella ataxia	74	88	-14	yes
7	133	Oligodendroma (grade 2)	87	74	13	no
8	26	Spinal cord infarct	92	95	-3	no
9	63	Multiple Sclerosis	33	41	-8	no
10	76	Multiple Sclerosis	75	79	-4	no
11	23	Multiple Sclerosis	93	95	-2	no
12	23	Spinal muscular atrophy	95	97	-2	no
13	87	Stroke haemorrhage	25	26	-1	no
14	17	Cervical myelopathy	94	94	0	no
15	18	Parkinson	93	94	-1	no
16	27	Cavernous haemangioma	80	81	-1	no
17	38	Multiple Sclerosis	84	95	-11	no
18	41	Multiple Sclerosis	84	89	-5	no