Comparison of rehabilitation outcomes for long term neurological conditions: a cohort analysis of the Australian rehabilitation outcomes centre dataset for adults of working age

Lynne Turner-Stokes  
*King's College London*

Roxana Vanderstay  
*King's College London*

Tara L. Stevermuer  
*University of Wollongong, hurst@uow.edu.au*

Frances D. Simmonds  
*University of Wollongong, francess@uow.edu.au*

Fary Khan  
*University of Melbourne*

*See next page for additional authors*
Comparison of rehabilitation outcomes for long term neurological conditions: a cohort analysis of the Australian rehabilitation outcomes centre dataset for adults of working age

Abstract

Objective

To describe and compare outcomes from in-patient rehabilitation (IPR) in working-aged adults across different groups of long-term neurological conditions, as defined by the UK National Service Framework.

Design

Analysis of a large Australian prospectively collected dataset for completed IPR episodes (n = 28,596) from 2003-2012.

Methods

De-identified data for adults (16–65 years) with specified neurological impairment codes were extracted, cleaned and divided into ‘Sudden-onset’ conditions: (Stroke (n = 12527), brain injury (n = 7565), spinal cord injury (SCI) (n = 3753), Guillain-Barré syndrome (GBS) (n = 805)) and ‘Progressive/stable’ conditions (Progressive (n = 3750) and Cerebral palsy (n = 196)). Key outcomes included Functional Independence Measure (FIM) scores, length of stay (LOS), and discharge destination.

Results

Mean LOS ranged from 21–57 days with significant group differences in gender, source of admission and discharge destination. All six groups showed significant change (p<0.001) between admission and discharge that was likely to be clinically important across a range of items. Significant between-group differences were observed for FIM Motor and Cognitive change scores (Kruskal-Wallis p<0.001), and item-by-item analysis confirmed distinct patterns for each of the six groups. SCI and GBS patients were generally at the ceiling of the cognitive subscale. The ‘Progressive/stable’ conditions made smaller improvements in FIM score than the ‘Sudden-onset conditions’, but also had shorter LOS.

Conclusion

All groups made gains in independence during admission, although pattern of change varied between conditions, and ceiling effects were observed in the FIM-cognitive subscale. Relative cost-efficiency between groups can only be indirectly inferred. Limitations of the current dataset are discussed, together with opportunities for expansion and further development.

Keywords

working, term, long, outcomes, rehabilitation, comparison, adults, dataset, centre, australian, analysis, cohort, conditions, age, neurological
Publication Details

Authors
Lynne Turner-Stokes, Roxana Vanderstay, Tara L. Stevermuer, Frances D. Simmonds, Fary Khan, and Kathy Eagar

This journal article is available at Research Online: http://ro.uow.edu.au/ahsri/650
Comparison of Rehabilitation Outcomes for Long Term Neurological Conditions: A Cohort Analysis of the Australian Rehabilitation Outcomes Centre Dataset for Adults of Working Age

Lynne Turner-Stokes¹ *, Roxana Vanderstay¹, Tara Stevermuer², Frances Simmonds², Fary Khan³, Kathy Eagar²

¹ King's College London, School of Medicine, Department of Palliative Care, Policy and Rehabilitation, London, United Kingdom, ² University of Wollongong, Australasian Rehabilitation Outcomes Centre / Australian Health Services Research Institute, Wollongong, Australia, ³ University of Melbourne and Royal Melbourne Hospital, Department of Rehabilitation, Melbourne, Australia

* lynne.turner-stokes@dial.pipex.com

Abstract

Objective
To describe and compare outcomes from in-patient rehabilitation (IPR) in working-aged adults across different groups of long-term neurological conditions, as defined by the UK National Service Framework.

Design
Analysis of a large Australian prospectively collected dataset for completed IPR episodes (n = 28,596) from 2003-2012.

Methods
De-identified data for adults (16–65 years) with specified neurological impairment codes were extracted, cleaned and divided into ‘Sudden-onset’ conditions: (Stroke (n = 12,527), brain injury (n = 7,565), spinal cord injury (SCI) (n = 3,753), Guillain-Barré syndrome (GBS) (n = 805)) and ‘Progressive/stable’ conditions (Progressive (n = 3,750) and Cerebral palsy (n = 196)). Key outcomes included Functional Independence Measure (FIM) scores, length of stay (LOS), and discharge destination.

Results
Mean LOS ranged from 21–57 days with significant group differences in gender, source of admission and discharge destination. All six groups showed significant change (p<0.001) between admission and discharge that was likely to be clinically important across a range of items. Significant between-group differences were observed for FIM Motor and Cognitive...
change scores (Kruskal-Wallis p<0.001), and item-by-item analysis confirmed distinct patterns for each of the six groups. SCI and GBS patients were generally at the ceiling of the cognitive subscale. The ‘Progressive/stable’ conditions made smaller improvements in FIM score than the ‘Sudden-onset conditions’, but also had shorter LOS.

Conclusion

All groups made gains in independence during admission, although pattern of change varied between conditions, and ceiling effects were observed in the FIM-cognitive subscale. Relative cost-efficiency between groups can only be indirectly inferred. Limitations of the current dataset are discussed, together with opportunities for expansion and further development.

Introduction

In 2005, the UK Department of Health published a National Service Framework (NSF) for long term neurological conditions (LTNC)[1]. Previous NSFs had already focused specifically on Older Adults and on Children, so the primary focus of the NSF for LTNC was on adults of working age (predominantly 16–65 years). Because of the diversity of presentation and the many diagnoses covered by this term, the NSF took a novel approach to the classification of neurological conditions, grouping them by pattern of presentation as follows:

• Sudden-onset conditions e.g. stroke, brain and spinal cord injury, acute polyneuropathies (e.g. Guillain-Barré syndrome (GBS)).

• Progressive conditions e.g. multiple sclerosis (MS), Parkinson’s Disease (PD), motor neurone disease (MND), chronic demyelinating polyneuropathies.

• Stable conditions (with changing needs due to development or ageing), e.g. cerebral palsy (CP), post-polio.

Although rehabilitation outcomes are well described in the literature for the Sudden-onset category—particularly for stroke and traumatic brain injury[2, 3], they are less well described for other neurological conditions (such as PD, neuropathies or cerebral palsy). One reason for this is the small number of patients within each diagnostic group. By combining patients with similar presentations into one larger group, it may be possible to explore outcomes and draw conclusions on a broader evidence base than can be achieved for each condition individually. Exactly how conditions should be grouped within those broad categories, however, remains open to question.

The analysis of prospectively-collected datasets provides an important opportunity to evaluate and compare outcomes across different conditions. Although cohort analyses do not provide direct evidence of the effectiveness of rehabilitation, they can afford more detailed information about which types of patients benefit from which types of treatment and in what ways[4, 5]. Importantly, they furnish generalisable information about the changes that occur in the course of real-life clinical practice (practice-based evidence), which is of interest to providers and purchasers of rehabilitation services[6]. On the other hand, the findings must be interpreted with a degree of caution where there are less rigorous standards for data collection, or in health settings where reimbursement is dependent on the demonstration of functional gain.
In Australia, there is no direct link between outcome and payment for rehabilitation services. The Australasian Rehabilitation Outcomes Centre (AROC) holds a large centralised database, which gathers a standard set of information on both process and outcomes for every person admitted for inpatient rehabilitation[7]. Established in 2002 as a joint initiative of the Australian rehabilitation sector (providers, payers, regulators and consumers), the dataset comprises case episode data for admissions for rehabilitation from participating services across Australia and New Zealand (currently almost 950,000 episodes of care from 266 facilities). The database provides a national benchmarking service as well as providing information to improve understanding of factors that influence quality of care and patient’s rehabilitation outcomes.

In the UK, an equivalent national dataset for specialist neurorehabilitation has been developed through the UK Rehabilitation Outcomes Collaborative (UKROC). The UKROC dataset represents the In-patient Rehabilitation module of the Long Term Neurological Conditions Dataset[8]. The design is modelled closely on the AROC dataset, but extends it in some areas. As the database is still in development, there is opportunity to learn from analyses of other large datasets to determine what further information may need to be collected alongside the core data, in order to address the critical questions in neurological rehabilitation over the coming decade. Crucial to the success of clinical datasets, however, is the engagement of clinicians to ensure that data are as complete and as accurate as possible. They need a frame of reference against which to compare their experience, and to gauge their outcomes in treating not only for the common conditions, but also for the rarer ones.

The primary objective of this analysis is to describe and compare outcomes for in-patient rehabilitation across a wide range of long-term neurological conditions in working-aged adults, which is the predominant focus for many of the more specialised rehabilitation services (Level 1 and 2) contributing to the UKROC database[9].

- We wished to explore how episode data might reasonably be collated into groups according to the NSF definitions for future analyses, and to present the information in a form that is meaningful to clinicians.

- We also wished to determine the extent to which the AROC dataset could be used to distinguish and describe these different groups.

The results will also assist with the identification of any additional information or approaches that should be included in future versions of the Australasian and UK datasets to align them more closely.

Materials and Methods

The AROC dataset

The full AROC dataset includes 42 items: socio-demographic, medical (impairment codes, co-morbidities, complications), episode items (admission dates), funding and employment details, and outcome data (patient level of function at admission and discharge)[10]. Within the AROC dataset, there are four principal 'Impairment categories' for neurological conditions. Each category is further subdivided into specified 'Impairment codes' (see Table 1).

The primary outcome measure is the Functional Independence Measure (FIM)[11]. The FIM is an 18-item global measure of independence in activities of daily living, subdivided into two subscales: FIM-Motor (13 items covering self-care, sphincters, transfers and locomotion) and FIM-Cognitive (5 items covering communication and social cognition). Each item is scored on a range of 1 (total dependence) to 7 (full independence). AROC holds a territory licence for use of the FIM (a trademark of the Uniform Data System for Medical Rehabilitation,
Clinical staff are required to complete FIM training and to sit a credentialing exam every 2 years.

All data received by AROC are screened for errors and missing data and, if necessary, the submitting facility is requested to review and correct any inconsistencies.

### Data extraction

For this analysis, we used data gathered from Australian facilities (n = 142) for episodes discharged in the years 2003 to 2012. De-identified data for inpatient rehabilitation (IPR) episodes...
for adults aged 16–65 years with a specified AROC Impairment Code within the four neurological impairment categories were extracted from the AROC data collection and transferred to SPSS version 18.0 for Windows for analysis. A total of $n = 36,672$ case episodes were recorded within the period.

For the purposes of our analysis, we were interested in comparing outcomes across the different conditions for those who completed a rehabilitation programme. The first step was therefore to classify episodes into ‘complete’ or ‘incomplete’. Case episodes were considered complete if they met all three of the following criteria:

- Discharged to usual or interim accommodation, or on other non-acute/sub-acute settings (i.e. incomplete episodes were patients who died ($n = 68$), self-discharged ($n = 361$), or were transferred back to the acute hospital ($n = 4702$) or missing data ($n = 715$)).
- Length of stay (LOS) was between 5–500 days (i.e. incomplete episodes were short admissions for assessment only ($n = 1150$), very long-stay episodes ($n = 13$) and missing LOS ($n = 54$)).
- A valid FIM score (i.e. case episodes with missing or invalid FIM scores ($n = 1013$) were incomplete).

Based on these criteria, the dataset was divided into 28,596 complete and 8076 incomplete episodes as shown in the Fig 1. Table 1 shows the categorisation of these included case episodes by AROC Impairment code. The profile was broadly similar although the incomplete group included a slightly higher proportion of traumatic brain and spinal cord injuries.

**Grouping of condition according to NSF categories**

We first explored the extent to which episode data for the different neurological conditions could be collated into groups according to the NSF definitions. In an initial stage exploration, we separated the dataset for complete episodes into 12 condition categories as shown in

---

**Fig 1. Flow chart—data exclusion.** The figure shows the different stages of the data cleaning process to obtain a dataset comprising completed episodes of rehabilitation with valid FIM scores. The most common reason for exclusion, accounting for more than half the excluded cases, was transfer back to the acute hospital setting for medical or surgical management. Many such cases will have returned to rehabilitation in a separate episode. Multiple episodes for the same individual were not linked in this analysis, but the process of concatenation will support linkage of serial episodes in future analyses.

doi:10.1371/journal.pone.0132275.g001
Table 2. CP was the only identifiable condition within the ‘Stable’ category (there is no AROC impairment code for Post-polio). The 12 categories were then compared to determine whether similarities between them would support further collation under the three main NSF groupings. In addition to analysis of between-group differences in FIM total and subscale score, we examined visually the graphic presentations of the disability profile (FIM-Splats) for each of the 12 categories[13]. The FIM-Splat is a radar chart showing the median scores on admission and discharge for each of the 18 FIM items. Three assessors independently examined the 12 FIM-Splats, grouped them on the basis of observed similarities, and then conferred to reach consensus. Six main groups were identified for the final analysis (see Fig 2). These were collated into two broad NSF categories (‘Sudden onset’ and ‘Progressive /stable’ conditions) due to the small size of the Stable group.

Outcomes of interest

The primary outcomes of interest were change in patient functional status, hospital length of stay (LOS) and discharge destination. In a more detailed analysis, we explored the specific items of function within the FIM (self-care, bladder/ bowel continence, mobility, cognition etc.) that did and did not change within each of the 6 groups.

Statistical Analysis

There is continued debate over the appropriate methods for statistical analysis, in view of concerns about excessive mathematical manipulation of ordinal data[14]. Rasch analysis offers the theoretical option of creating interval level data from ordinal scales, but as yet there is a lack of consensus in how to achieve this. Rasch models are not yet available for every condition, and neither are bedside computers to provide instantaneous interval-level conversion. In routine practice, clinicians must interpret the ordinal data as best they can. To make this analysis meaningful for them, we have taken a simple pragmatic approach to this analysis using standard descriptive and non-parametric methods.

Descriptive analysis included the counts, percentage, mean, standard deviation, median and inter-quartile range as appropriate for demographic, LOS and discharge destination (percent

Table 2. Grouping for analysis.

<table>
<thead>
<tr>
<th>NSF Category</th>
<th>12 groups</th>
<th>N =</th>
<th>6 groups</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sudden-onset</td>
<td>Cerebrovascular Accident</td>
<td>12527</td>
<td>Stroke</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subarachnoid haemorrhage</td>
<td>1014</td>
<td>Other acquired brain injury (ABI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Brain injury—traumatic</td>
<td>4351</td>
<td>Other acquired brain injury (ABI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Brain injury—non-traumatic</td>
<td>2200</td>
<td>Other acquired brain injury (ABI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spinal cord injury—traumatic</td>
<td>1677</td>
<td>Spinal cord injury (SCI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spinal cord injury—non-traumatic</td>
<td>2076</td>
<td>Spinal cord injury (SCI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Guillain-Barré Syndrome</td>
<td>805</td>
<td>Guillain-Barré Syndrome (GBS)</td>
<td></td>
</tr>
<tr>
<td>Progressive</td>
<td>Multiple sclerosis</td>
<td>2292</td>
<td>Progressive</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parkinson’s disease</td>
<td>847</td>
<td>Progressive</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neuromuscular disease</td>
<td>408</td>
<td>Progressive</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other (chronic) polyneuropathies</td>
<td>203</td>
<td>Progressive</td>
<td></td>
</tr>
<tr>
<td>Stable</td>
<td>Cerebral Palsy</td>
<td>196</td>
<td>Cerebral Palsy (CP)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>28596</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
discharged to community/remaining in hospital system) collated by condition. As age and LOS are interval data, between group comparisons were determined by one-way unrelated ANOVA with post hoc Bonferroni correction.

Previous studies have reported mean FIM gain and FIM efficiency (FIM gain/ length of stay)[2, 3], so these figures are given for the sake of comparison. However, as the FIM is an ordinal scale, we preferred the more conservative approach of using non-parametric statistics for analysis of FIM data[15].

Fig 2. Radar Charts (FIM-Splats) showing median FIM scores on admission and discharge for the 12 groups. The FIM-Splat provides graphic presentation of the disability profile in a radar chart. The 18 items are arranged as ‘spokes of the wheel’ and the Levels from 1 (total dependence) to 7 (total independence) run from the centre outwards. Thus a perfect score would be demonstrated as a large circle. The group median scores for each item are plotted for admission and discharge. The difference between median scores on admission and discharge is depicted by the shaded area.

doi:10.1371/journal.pone.0132275.g002
Spearman Rank tests were used to explore the association between Admission FIM scores and LOS.

Within group comparisons of functional status (FIM) on admission and discharge were tested by Wilcoxon signed rank tests, for individual items, as well as subscales and total scores.

Overall between-group comparisons were determined by Kruskal-Wallis tests. Post hoc group-by-group comparisons were undertaken using Mann Whitney tests. P values were multiplied by the number of tests to correct for multiple comparisons. Corrected p values of <0.05 were considered significant.

Ethics Statement
This study was approved by the Human Research and Ethics Committee, The Royal Melbourne Hospital (HREC 2010.004)

Results
Because nearly 22% of the sample comprised ‘incomplete’ episodes, it was pertinent to examine the characteristics of these cases, which are presented in Table 3. Overall the incomplete episode cases were younger (mean difference 2.2 years p<0.0001) than the completed group, with a greater proportion of males (64.9 vs. 60.8, Chi squared 41.7 p<0.0001) and had a longer length of stay (mean difference 3.8 days, p<0.0001). Over half (57.8%) were transferred back to acute hospital settings, 4.5% self-discharged at their own risk and <1% died during their IPR admission. Unsurprisingly, the majority of incomplete episodes were in the sudden onset conditions, particularly brain and spinal cord injury, where a higher incidence of serious inter-current illness is expected. The relatively high proportion of patients returning to the acute sector in this sample may reflect the relatively early transition to rehabilitation following injury in Australia and also the lack of investigation facilities in many free-standing specialist rehabilitation units necessitating a formal discharge each time a patient is referred to another hospital for investigation or treatment.

The remainder of the analysis includes only case episodes for a completed programme of rehabilitation (n = 28,296), divided into the six main groups. Demographics are shown in Table 4. The large majority of stroke, brain injury and GBS patients (>90%) were admitted from acute hospital services, with less than 10% coming from home. By contrast over 40% of patients with Progressive conditions and CP were admitted from home. In the spinal cord injury group, over three quarters came from hospital, but the rest were admitted from home or other non-hospital settings. Similarly, the large majority of patients were discharged back to their usual accommodation at the end of their rehabilitation programme (90.1% overall), although a significant proportion of sudden onset conditions (8.4% overall) were discharged to interim and other accommodation, compared with 4% in the progressive and stable conditions.

Gender and age distribution were largely as expected—the predominance of females in the Progressive group being largely due to the preponderance of MS patients. All groups spanned at least the age range 17–65 years. One-way ANOVA tests confirmed significant between group differences in age (p<0.001). Post hoc analysis with Bonferroni correction confirmed significant differences between all groups except ABI and CP (p = 1.000).
### Table 3. Characteristics of the incomplete episodes (n = 8076).

<table>
<thead>
<tr>
<th></th>
<th>Stroke</th>
<th>ABI</th>
<th>SCI</th>
<th>GBS</th>
<th>Progressive</th>
<th>CP</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Episodes(%)</td>
<td>2934 (36.3)</td>
<td>2618 (32.4)</td>
<td>1669 (20.7)</td>
<td>142 (1.8)</td>
<td>661 (8.2)</td>
<td>52 (0.6)</td>
<td>8076 (100.0)</td>
</tr>
<tr>
<td>Male N (%)</td>
<td>1845 (62.9)</td>
<td>1718 (68.0)</td>
<td>1223 (73.3)</td>
<td>83 (58.5)</td>
<td>287 (43.4)</td>
<td>23 (44.2)</td>
<td>5242 (64.9)</td>
</tr>
<tr>
<td>Age (Mean (SD))</td>
<td>53.6 (10.2)</td>
<td>41.2 (14.6)</td>
<td>43.1 (14.9)</td>
<td>48.7 (13.7)</td>
<td>50.5 (10.6)</td>
<td>35.0 (12.8)</td>
<td>46.9 (14.1)</td>
</tr>
<tr>
<td>Length of stay, days (Mean (SD))</td>
<td>32.6 (42.1)</td>
<td>45.7 (72.1)</td>
<td>54.1 (64.3)</td>
<td>35.6 (61.7)</td>
<td>21.4 (31.2)</td>
<td>16.4 (18.9)</td>
<td>40.3 (58.6)</td>
</tr>
</tbody>
</table>

#### Discharge destination N(%)
- Home/community: 562 (19.2) 511 (19.5) 150 (9.0) 40 (28.2) 202 (30.6) 27 (51.9) 1492 (18.5)
- Acute Hospital: 1800 (61.3) 1411 (53.9) 1035 (62.0) 75 (52.8) 323 (48.9) 20 (38.5) 4664 (57.8)
- Own risk: 168 (5.7) 123 (4.7) 25 (1.5) 8 (5.6) 37 (5.6) 0 (0.0) 361 (4.5)
- Care / other: 198 (6.7) 288 (11.0) 240 (14.4) 8 (5.6) 38 (5.7) 4 (7.7) 776 (9.6)
- Died: 31 (1.1) 18 (0.7) 8 (0.5) 1 (0.7) 10 (1.5) 0 (0.0) 68 (0.8)
- Missing: 175 (6.0) 267 (10.2) 211 (9.2) 10 (7.0) 51 (7.7) 1 (1.9) 715 (8.9)

**do:**10.1371/journal.pone.0132275.t003

### Table 4. Demographics within the six groups for the complete episodes (n = 28,596).

<table>
<thead>
<tr>
<th></th>
<th>Stroke</th>
<th>ABI</th>
<th>SCI</th>
<th>GBS</th>
<th>Progressive</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Episodes</td>
<td>12527</td>
<td>7565</td>
<td>3753</td>
<td>805</td>
<td>3750</td>
<td>196</td>
</tr>
<tr>
<td>Admitted from:</td>
<td>Missing data n = 439 (1.5%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home/community</td>
<td>1097 (8.9)</td>
<td>526 (7.0)</td>
<td>541 (14.7)</td>
<td>44 (5.5)</td>
<td>1453 (40.0)</td>
<td>98 (50.0)</td>
</tr>
<tr>
<td>Hospital</td>
<td>11077 (90.0)</td>
<td>6882 (91.3)</td>
<td>2887 (78.4)</td>
<td>738 (92.6)</td>
<td>2142 (59.0)</td>
<td>97 (49.5)</td>
</tr>
<tr>
<td>Other (eg care)</td>
<td>137 (1.1)</td>
<td>131 (1.7)</td>
<td>255 (6.9)</td>
<td>15 (1.9)</td>
<td>36 (1.0)</td>
<td>1 (0.5)</td>
</tr>
<tr>
<td>Discharge to:</td>
<td>No missing data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usual accommodation</td>
<td>11401 (91.0)</td>
<td>6607 (87.3)</td>
<td>3277 (87.3)</td>
<td>751 (93.3)</td>
<td>3541 (94.4)</td>
<td>188 (95.9)</td>
</tr>
<tr>
<td>Interim accommodation</td>
<td>824 (6.6)</td>
<td>822 (10.9)</td>
<td>365 (9.7)</td>
<td>48 (6.0)</td>
<td>155 (4.1)</td>
<td>4 (2.0)</td>
</tr>
<tr>
<td>Other</td>
<td>302 (2.4)</td>
<td>136 (1.8)</td>
<td>111 (3.0)</td>
<td>6 (0.7)</td>
<td>54 (1.4)</td>
<td>4 (2.0)</td>
</tr>
<tr>
<td>Gender (%)</td>
<td>Male 61.9 67.5 69.0 55.9 38.2 43.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Female 38.1 32.5 31.0 44.1 61.8 56.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age Mean (St Dev)</td>
<td>54.4 (9.6)</td>
<td>42.6 (14.9)</td>
<td>44 (14.5)</td>
<td>47.1 (13.4)</td>
<td>51.4 (10.6)</td>
<td>40.6 (13.8)</td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>57 (50–62)</td>
<td>45 (29–56)</td>
<td>46 (32–57)</td>
<td>50 (38–59)</td>
<td>54 (45–60)</td>
<td>41 (29–52)</td>
</tr>
<tr>
<td>Length of stay (LOS), days Mean (St Dev)</td>
<td>34.6 (31.4)</td>
<td>36.2 (45.2)</td>
<td>56.9 (59.5)</td>
<td>35.1 (36.2)</td>
<td>22.9 (21.5)</td>
<td>21.1 (12.8)</td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>25 (14–45)</td>
<td>22 (13–41)</td>
<td>36 (17–75)</td>
<td>22 (14–42)</td>
<td>17 (12–26)</td>
<td>18 (12–26)</td>
</tr>
<tr>
<td>Range</td>
<td>5–480 5–499 5–494 5–313 5–419 5–93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spearman correlation LOS vs Total FIM score on admission</td>
<td>-0.655 (p&lt;0.001)</td>
<td>-0.627 (p&lt;0.001)</td>
<td>-0.559 (p&lt;0.001)</td>
<td>-0.667 (p&lt;0.001)</td>
<td>-0.410 (p&lt;0.001)</td>
<td>-0.255 (p&lt;0.001)</td>
</tr>
<tr>
<td>FIM gain</td>
<td>No missing data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIM-Motor gain (Mean + 95%CI)</td>
<td>21.6 (20.3–21.9)</td>
<td>19.9 (19.5–20.4)</td>
<td>20.2 (19.6–20.8)</td>
<td>26.9 (25.7–28.3)</td>
<td>11.8 (11.4–12.2)</td>
<td>9.5 (8.2–10.9)</td>
</tr>
<tr>
<td>FIM-Cognitive gain (Mean + 95%CI)</td>
<td>3.5 (3.4–3.6)</td>
<td>5.5 (5.4–5.7)</td>
<td>0.6 (0.5–0.7)</td>
<td>0.9 (0.8–1.1)</td>
<td>0.9 (0.8–1.1)</td>
<td>0.9 (0.6–1.3)</td>
</tr>
<tr>
<td>FIM-Total gain (Mean + 95%CI)</td>
<td>25.1 (24.8–25.5)</td>
<td>25.5 (24.9–25.9)</td>
<td>20.8 (20.2–21.4)</td>
<td>27.9 (26.6–29.3)</td>
<td>12.8 (12.4–13.2)</td>
<td>10.5 (9.0–11.9)</td>
</tr>
<tr>
<td>FIM Efficiency (FIM total gain/LOS)</td>
<td>0.73</td>
<td>0.70</td>
<td>0.36</td>
<td>0.79</td>
<td>0.56</td>
<td>0.50</td>
</tr>
</tbody>
</table>

IQR = inter-quartile range; CI = Confidence interval; FIM = Functional Independence Measure.

**do:**10.1371/journal.pone.0132275.t004
Mean length of stay ranged from 21 to 57 days. One-way ANOVA tests confirmed significant between group differences in length of stay (p < 0.001). Post hoc analysis with Bonferroni correction confirmed three distinct groups:

- SCI (Longer stay) (typically mean 8 weeks)
- Stroke, ABI and GBS (Medium stay) (typically mean 5 weeks)
- Progressive and CP (Shorter stay) (typically mean 3 weeks)

As expected there was a strong negative correlation between LOS and FIM total score on admission, but correlations were stronger for the four 'Sudden-onset' conditions (rho -0.56 to -0.66) than for the 'Progressive/Stable' conditions (rho -0.26 to -0.41). Gains were smaller for the progressive and sudden onset groups, but were proportionate to their shorter lengths of stay.

**Within groups analysis**

Table 5 shows the median (IQR) for FIM total and subscale scores on admission and discharge, together with the results of Wilcoxon tests. All six groups showed statistically significant change (p < 0.001) between admission and discharge in both motor and cognitive subscales as well as total scores.

**Between groups analysis**

Fig 3 illustrates the FIM motor and cognitive change scores between admission and discharge for the six groups. Kruskal-Wallis tests confirmed significant overall between-group differences (p < 0.001) for both FIM-Motor and FIM-Cognitive change scores. Post hoc Mann Whitney tests for change in Motor function confirmed significant differences between all groups (corrected p < 0.01), except between ABI and SCI (p = 1.000). Post hoc tests for change in Cognitive

---

**Table 5. Median (IQR) FIM total and subscale scores on admission and discharge, with Wilcoxon Signed Rank tests for change between admission and discharge.**

<table>
<thead>
<tr>
<th></th>
<th>Stroke</th>
<th>ABI</th>
<th>SCI</th>
<th>GBS</th>
<th>Progressive</th>
<th>CP</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Admission Scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Discharge Scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Change Scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIM Motor z score</td>
<td>198–332z = -93.6</td>
<td>155–292z = -70.4</td>
<td>155–332z = -49.3</td>
<td>2411–412z = -24.2</td>
<td>94–172z = -49.0</td>
<td>71–142z = -10.9</td>
<td>166–302z = -138.9</td>
</tr>
<tr>
<td>FIM Cognitive z score</td>
<td>20–52z = -74.1</td>
<td>40–92z = -63.5</td>
<td>00–02z = -17.0</td>
<td>00–02z = -12.3</td>
<td>00–11z = -25.4</td>
<td>00–11z = -6.2</td>
<td>10–52z = -103.7</td>
</tr>
<tr>
<td>FIM Total z score</td>
<td>2210–372z = -94.6</td>
<td>208–372z = -72.7</td>
<td>165–342z = -49.4</td>
<td>2412–422z = -24.3</td>
<td>104–192z = -49.2</td>
<td>82–152z = -11.0</td>
<td>188–342z = -140.9</td>
</tr>
</tbody>
</table>

*All significant at p < 0.001.

DOI: 10.1371/journal.pone.0132275.t005
function similarly confirmed significant differences between all groups except between GBS, progressive conditions and CP (p = 1.000).

Item by item analysis
All FIM items showed statistically significant change in all conditions (Wilcoxon p < 0.001). However the size of the change was often small. Table 6 shows the analysis of group level data and records the improvement (between admission and discharge) in median FIM score for each item in each of the six groups.

- In the FIM motor items, all groups showed improvement of 2 or more points across at least two items. The Progressive group showed smaller changes, but nevertheless change was seen across all motor items except sphincters, despite the relatively short lengths of stay.
- In the FIM cognitive items, the stroke and ABI groups showed improvements of 0–2 points, but there was no improvement for SCI or GBS.

Fig 4 shows the radar charts (‘FIM-Splats’) of the median item scores on admission and discharge for the six groups. The different pattern for each group is clearly seen. In SCI and GBS, the lack of change in cognitive scores reflects a ceiling effect at admission. Near-ceiling effects also account for the small degree of change in Progressive conditions. In CP, there were mild cognitive deficits at baseline, but these largely remained static.

Discussion
In this analysis of a large Australian dataset, we have compared outcomes from in-patient rehabilitation across groups of long term neurological conditions, categorised according to the UK NSF for Long Term Neurological Conditions into Sudden-onset, Progressive and Stable Conditions. The analysis was centred on adults of working age to reflect the emphasis of the NSF, and also the predominant focus of many of the more specialised rehabilitation services in the UK[16].

The literature contains a number of other analyses from large multi-centre rehabilitation datasets—notably held by Uniform Data Systems in the USA which is undoubtedly the largest in the world. Previous reports (including other published analyses of the AROC dataset) have tended to focus on single conditions, such as stroke[2, 17, 18], traumatic brain injury[3], spinal cord injury[19, 20] multiple sclerosis[21, 22] or Guillain-Barré Syndrome[23, 24]. These have either provided a general description for benchmarking information[2, 3, 22, 23] or compare outcomes across different rehabilitation settings[17] or for different racial and ethnic groups [18]. Ottenbacher et al (2004) examined year-on-year trends in length of stay, living setting, functional outcome, and mortality[25]. Several authors have used Rasch analysis to examine differential item functioning (order of difficulty for individual items) across different neurological conditions[26, 27] and smaller single centre analyses from the UK have examined changes in FIM for a general neurorehabilitation sample[28, 29]. However, this is one of very few large clinical dataset analyses from outside the US and the first to compare functional outcomes at item level across different neurological conditions grouped according to the NSF categories.

The data were collected in the course of routine clinical practice and we have deliberately kept the analysis simple, so that clinicians can interpret it and use it to compare their own practice. FIM-splats have proven popular in clinical settings, both in the UK and in Australia, for providing an ‘at-a-glance’ impression of the areas in which change has occurred. The use of FIM-splats to inform clinical grouping based on the profile of change across individual FIM items is a novel approach, which we believe will provide a useful basis for future analyses of the
dataset. For example, small groups of rarer conditions may be included within the group with the closest-matching FIM profile. Alternatively, where the condition is to be considered separately, examination of the FIM profile may be used to inform the selection of an appropriate comparator group. This approach also may have potential application for other large datasets around the world.

Although the AROC dataset does not directly provide categorisation into the three main NSF groups, it was possible to group the conditions by impairment code. The logic of separating groups into the NSF categories was to some extent borne out by the analysis. At the crudest

---

**Fig 3. FIM Motor and Cognitive change scores.** Medians are denoted by solid black lines while the top and bottom box edges denote the first and third quartile. The T-bars or whiskers denote the largest and smallest data within 1.5 times the interquartile range. The small circles are outliers, values that do not fall in the inner fences. The extreme values amongst outliers are marked with an asterisk. These are cases that have values more than three times the height of the boxes.
level there were clear differences between the 'Sudden-onset' conditions and the 'Progressive' and 'Stable' conditions in terms of the source of admission, length of stay, discharge destination and functional gain. However to group the conditions into just two or three main categories would miss important differences between them. Between groups analysis, together with examination of the different patterns of improvement as shown in the FIM-Splats, suggests that the six-group analysis performed here represents an appropriate balance between capturing clinically important differences and maintaining a manageable number of groups for statistical comparison.

Within groups analysis showed that all groups made statistically significant changes in FIM score between admission and discharge, at the levels of total, subscale, and individual item scores.

- All six groups showed substantial changes in motor function. Even at item-level, all groups made gains that were likely to be clinically important (see below) across a wide range of items.

- With respect to cognitive and communicative function, Stroke and ABI patients showed moderate change, and the Progressive and CP groups showed smaller changes. However, the majority SCI and GBS patients were already at the upper limit of the scales on admission.

Across all domains and all conditions the scores may go either up or down between admission and discharge, as illustrated in Fig 2. Even where the neurological condition does not directly affect the brain (i.e. spinal cord injury, GBS), a proportion of patients do have problems in the cognitive domains (for example due to inter-current infection, metabolic disturbance or occult brain pathology), which have a general impact on function and are addressed during the

<table>
<thead>
<tr>
<th>FIM Items</th>
<th>Stroke</th>
<th>ABI</th>
<th>SCI</th>
<th>GBS</th>
<th>Progressive</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eating</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Grooming</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Bathing</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Dress-up</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dress-low</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Toilet</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Bladder</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bowel</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bed transfer</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Toilet transfer</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Tub transfer</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Walk</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Stairs</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td><strong>Cognitive</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Express</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Social</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Problem</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Memory</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Change in median value: in bold >2.

doi:10.1371/journal.pone.0132275.t006
rehabilitation process. Therefore, as other authors have highlighted[20, 30], the absence of change in FIM cognitive score may represent a ceiling effect of the scale itself, rather than a genuine lack of change. A number of solutions have been proposed, including the addition of further items to address cognitive / psychosocial function, to form the Functional Assessment Scale (FIM+FAM)[31, 32] Whilst the FIM+FAM may not significantly extend the scaling range of the FIM, there is evidence that it provides extended coverage of individual goals for rehabilitation on a qualitative level[33].

We recognise a number of specific limitations to this study.

1. As with analysis of any large dataset collected in the course of routine practice, there was significant attrition due to incomplete data. Although data are carefully checked and validated at the point of submission to AROC, and completeness of data entry is improving over time [22], the possibility of missing data, coding and reporting errors still exist, which could affect the results.

2. The AROC dataset records de-identified episodic data, resulting in the potential for more than one episode being reported against the one patient with all but one being incomplete. However, since 2013 calendar year AROC has introduced a new analysis practice called 'Concatenation'. Prior to outcomes analysis, AROC will identify groups of submitted episodes that can be joined to form a single AROC reporting episode, and consequently this will reduce the overall proportion of incomplete cases in the dataset.

Fig 4. FIM—Splat Motor and Cognitive change score. The FIM-Splat provides graphic presentation of the disability profile in a radar chart. The 18 items are arranged as ‘spokes of the wheel’ and the Levels from 1 (total dependence) to 7 (total independence) run from the centre outwards. Thus a perfect score would be demonstrated as a large circle. The group median scores for each item are plotted for admission and discharge. The difference between median scores on admission and discharge is depicted by the shaded area.

doi:10.1371/journal.pone.0132275.g004
3. The AROC dataset was not designed to separate episodes according to the NSF categories, and we recognise that the division between groups is not entirely clean. Similarly, as noted elsewhere[22], the AROC dataset does not distinguish the different patterns of onset of MS. The dataset did not provide robust enough information about the overall duration of the condition to separate these longer-standing patients with confidence, so cases were allocated to groups of the basis of the AROC impairment code alone. There are further opportunities for future analysis, for example to compare outcomes for patients admitted from acute services and from the community.

4. In such a large dataset even small differences are likely to reach statistical significance, even if they are of no clinical importance. A key challenge for this type of analysis is therefore to define what is meant by ‘clinically important’ change.

In terms of crude change in FIM score, the findings in this study are on par with other reports. For example, Beninato et al 2006[34] reported changes of 17, 3 and 22 respectively in FIM motor, cognitive and total scores in association with a Minimal Clinically Important Difference (MCID) in stroke. Our study showed changes of this order for the ‘Sudden onset’ conditions (see Tables 4 and 5). To our knowledge, MCID for FIM has not been reported for progressive conditions.

More importantly, however, the theoretical implication of improved independence following rehabilitation is that there should be a corresponding reduction in care needs, and therefore on-going costs in the community. The quantification of cost-benefits is a key challenge for any prospective data collection system. Granger and colleagues in the USA have reported a change of 1 point on the total FIM scale to equate to approximately 5 minutes of care per day for TBI patients[35], 3.32 minutes for stroke[36] and 3.38 minutes for MS[21]. Although it cannot be assumed that care costs translate across different health and social care cultures, were we to apply these estimates, the mean change in FIM recorded in this series (see Table 4) would equate to a reduction of approximately 14.8 care hours per week for TBI, 6.3 hours/week for stroke and 5.0 hrs per week for MS. However, this analysis is over simplistic, as Granger himself also points out. Some FIM items are more predictive of care requirement than others, and this may vary across the different conditions. For example, in MS locomotion and tub transfers were the strongest predictors[21], whereas for TBI, cognition and the need for support to maintain safety was a key factor[35]. At the very least, therefore, item level analysis is required to understand the impact of rehabilitation within the different conditions. Further analysis is on-going with this dataset to examine interval level changes and differential item functioning using Rasch and other techniques, and will be reported separately.

Implications for future data collection

Within the AROC dataset, the FIM serves a multifunctional role. The admission FIM score is applied as a casemix tool, the level of dependency being used as a proxy indication of need for rehabilitation and care. Change in FIM score from admission to discharge is reported as the primary outcome measure, and FIM efficiency (FIM gain / length of stay) is reported as a surrogate for service-efficiency. Although this model has the advantage of simplicity and minimising the burden of data collection, particularly in high throughput services, it may be too simple to provide adequate evaluation in the context of complex neurological rehabilitation. Data gathered in a tertiary neuro-rehabilitation setting in the UK demonstrate that, although the FIM correlates fairly well with needs for care and nursing, it is a poor predictor of needs for therapy and medical intervention[37]. Moreover, due to floor and ceiling effects, ‘FIM efficiency’ was shown not to be a sensitive indicator of cost-efficiency, other than in the middle
part of the score range[38]. There are also concerns about the validity of this type of mathematical manipulation of ordinal data[14].

Other measures are therefore required to provide a more complete evaluation of the complexity of ‘needs’ for rehabilitation as well as the ‘inputs’ (in particular staff resources) provided to meet them, before we can properly interpret measures of outcome and cost-efficiency. Over the last decade or so, newer tools have been developed and validated in the UK to provide more direct evaluation of these aspects. The Rehabilitation Complexity Scale[37] is a simple measure of rehabilitation needs. The Northwick Park nursing and therapy Dependency Scales [39–41] are measures of dependency which translate directly into estimates of staff time via a computerised algorithm, and have been used to provide a more direct evaluation of cost efficiency, especially for more dependent patients[38]. These are now incorporated into the UKROC dataset, as well as the option of recording the UK FAM items[32] to provide more comprehensive evaluation of cognitive and psychosocial outcome for those centres who wish to record them. The AROC dataset is also under review and the next iteration could potentially include a somewhat extended dataset to capture some of these parameters. Both centres are also exploring methods for pseudonymising patients, in order to track them through the system. The establishment of a common core of information for inclusion in these and other national rehabilitation datasets around the world would assist future international collaboration in outcomes analysis for rehabilitation.

Acknowledgments

We would like to thank all the centres and patients who contributed data to the AROC dataset, and all the AROC staff who are responsible for receiving, cleaning and collating the data to maintain this unique and high quality dataset. We are also grateful to the AROC Scientific and Clinical Advisory Committee for making the data available and allowing us the opportunity to undertake this analysis. Statistical advice was kindly provided by Paul Basset, Consultant Statistician.

Author Contributions

Conceived and designed the experiments: LTS. Analyzed the data: LTS RV TS. Contributed reagents/materials/analysis tools: LTS TS FS FK KE. Wrote the paper: LTS RV TS FS FK KE. Responsible for the overall drafting of the paper and manuscript preparation: LTS. Provided intellectual input with respect to clinical context: FK. Responsible for the development and ongoing integrity of the AROC database: KE FS. Responsibility as guarantor for the quality of the AROC dataset: KE. Responsible for the day-to-day running and management of the database: FS.

References